THE UNIVERSITY OF QUEENSLAND

AUSTRALIA

Structural studies of sorting nexin proteins from two distinct subfamilies

Blessy Abraham Paul

MPhil. Structural Chemistry

MSc. Biotechnology

BSc. Biotechnology

A thesis submitted for the degree of Doctor of Philosophy at

The University of Queensland in 2018

Institute for Molecular Bioscience

Abstract

Sorting nexin (SNX) proteins are a large family of proteins with critical roles in endocytosis, membrane trafficking and intracellular signalling. Each SNX protein contains a Phox homology (PX) domain that typically recognizes phosphoinositide lipids to enable their anchoring to defined organelles. SNXs also contain additional domains other than the PX domain thus allowing them to participate in a wide range of functions such as protein-protein interactions, membrane remodeling, lipid metabolism and other functions that are yet to be explored. My research focuses on the structural and functional characterization of two different sub-families of SNX proteins; those that have both PX and BAR (bin/amphiphysin/rvs) domains (SNX5, SNX6 and SNX32), and the poorly characterized SNX-RGS sub-family that contain a regulator of G-protein signaling (RGS) domain (SNX13, SNX14, SNX19 and SNX25). These protein families are required for membrane trafficking and lipid droplet formation, and are implicated in diseases including pathogen invasion and cerebellar ataxia respectively.

I first describe the structural mechanism for how a Chlamydial effector protein called IncE hijacks and recruits SNX5 related proteins to the bacterial inclusion membrane during Chlamydial infection (**Chapter 2)**. Using X-ray crystallography, I demonstrate that the C-terminal region of IncE forms a β-hairpin structure and binds to a unique α-helical insertion on the PX domain of SNX5, which is a separate site from the canonical PX lipid-binding site. Using isothermal titration calorimetry, I also investigate the binding mechanisms of the SNX5 homologs, SNX6 and SNX32 to IncE, and demonstrate that these three proteins share a common binding mechanism. These results suggest that the IncE protein mimics other SNX5-related proteins, where SNX5 PX might be functioning as a protein-binding scaffold that could potentially orchestrate the trafficking of certain transmembrane receptors.

The research presented in **Chapter 3** extends the work described in Chapter 2 confirming the IncE binding site in the SNX32 PX domain using X-ray crystallography. The crystal structure of SNX32 PX-IncE complex not only confirmed the biophysical data from chapter 2, but also revealed the binding mechanism to be identical to the SNX5PX-IncE binding. In addition, this is the first reported structure of the SNX32 protein in either an apo or ligand bound state. This result further strengthens the idea that SNX-BAR proteins can act as receptor recyclers in the cell.

In **Chapter 4** using biophysical and structural studies, I next investigated the potential functional role of the SNX-BAR proteins in transmembrane cargo trafficking. My biophysical experiments directly confirmed the recent indirect proteomic studies that

suggested cargos including CIMPR, IGF1R and SEMAC can bind to SNX-BAR proteins, and my first low-resolution crystal structure of the SNX5 PX domain with a transmembrane cargo, CIMPR, further confirms my hypothesis that endogenous human transmembrane cargos binds SNX5 related proteins in an analogous manner to IncE. Even though CIMPR does not possess sequence similarity with IncE, the identified hydrophobic patch in SNX5 is required and sufficient for the binding to the CIMPR peptide.

In **Chapter 5**, I explore the idea of synthesizing cyclic peptides based on my structure of the linear IncE peptide, which when bound to SNX-BAR proteins I showed forms a β-hairpin. The N- and C-termini of IncE were linked through various macrocyclisation approaches, and this yielded a cyclic IncE derived peptide that binds to SNX5 related SNX-BAR proteins with a nanomolar binding affinity that is 30x higher than the standard IncE peptide, and ~300x higher than cellular cargo molecules. I further demonstrate the ability of this cyclic peptide to inhibit the binding of a transmembrane cargo, CIMPR to SNX5 *in vitro*.

In **Chapter 6,** I describe my initial studies regarding whether the RGS domains of SNX-RGS proteins possess GTPase activating protein (GAP) activity for the small GTPase $G\alpha_s$. My studies reveal that RGS domains of SNX13, SNX14 and SNX25 binds to $G\alpha_s$, but do not possess GAP activity. My results lead to me to speculate that since SNX-RGS proteins bind to $G\alpha_s$ but do not regulate its GDP/GTP-bound activity states, these proteins may still act as negative regulators of $G\alpha_s$ signaling through ways that are yet to be explored.

In **Chapter 7** I provide a summary of my attempts to successfully express and purify the structurally uncharacterized PXA domain from the SNX-RGS subfamily, which is directly implicated in lipid droplet biogenesis. My trials have yielded soluble PXA protein and initial crystal hits, which provides a solid foundation for further structural and functional studies of this domain.

Finally, in **Chapter 8** I provide an overall summary of my work, both outcomes and future directions. Overall, my studies have revealed the crucial role of PX domains not only in membrane interactions but in direct protein-protein interactions required for cargo sorting, provides insights into how SNX proteins contribute to pathogen invasion and other diseases, and how structural information can be used to target these proteins with peptide tools in functional studies and in potential therapeutic approaches.

Declaration by author

This thesis is composed of my original work, and contains no material previously published or written by another person except where due reference has been made in the text. I have clearly stated the contribution by others to jointly-authored works that I have included in my thesis.

I have clearly stated the contribution of others to my thesis as a whole, including statistical assistance, survey design, data analysis, significant technical procedures, professional editorial advice, financial support and any other original research work used or reported in my thesis. The content of my thesis is the result of work I have carried out since the commencement of my higher degree by research candidature and does not include a substantial part of work that has been submitted to qualify for the award of any other degree or diploma in any university or other tertiary institution. I have clearly stated which parts of my thesis, if any, have been submitted to qualify for another award.

I acknowledge that an electronic copy of my thesis must be lodged with the University Library and, subject to the policy and procedures of The University of Queensland, the thesis be made available for research and study in accordance with the Copyright Act 1968 unless a period of embargo has been approved by the Dean of the Graduate School.

I acknowledge that copyright of all material contained in my thesis resides with the copyright holder(s) of that material. Where appropriate I have obtained copyright permission from the copyright holder to reproduce material in this thesis and have sought permission from co-authors for any jointly authored works included in the thesis.

Publications included in this thesis

Paul B, Kim HS, Kerr MC, Huston WM, Teasdale RD, Collins BM. (2017) Structural basis for the hijacking of endosomal sorting nexin proteins by *Chlamydia trachomatis*. *eLife* **6**: e22311

DOI: 10.7554/eLife.22311

Incorporated as Chapter 2

Contributor	Statement of contribution
Blessy Abraham Paul (Candidate)	Designed and performed experiments (70%) Wrote (70%) and edited the paper (10%) Responsible for Figs. 15A, 19A, 20A-C, 23A-E, 24A-C, 27A, Video 2
Hyun Sung Kim	Performed experiments (15%) Edited the paper (10%) Responsible for Figs. 16A-B, 25A-B, 26A-C
Marcus C Kerr	Performed experiments (10%) Edited the paper (10%) Responsible for Figs. 17, 18, Video 1
Wilhelmina M Huston	Performed experiments (5%) Edited the paper (10%) Responsible for Fig. 19C
Rohan D Teasdale	Edited the paper (10%)
Brett Collins	Wrote (30%) and edited the paper (50%) Responsible for Figs. 19B, 2, 22A-D, 27B-D

Submitted manuscripts included in this thesis

No manuscripts submitted for publication.

Other publications during candidature

Paul B, Collins BM. (2018) Molecular mechanism behind the interaction between SNX32 PX domain and Chlamydial effector protein IncE. (Draft revision for submission to *Protein Science*)

Simonetti B, Paul B, Weeratunga S, Teasdale RD, Collins BM, Cullen PJ. (2018) Molecular basis for the endosomal trafficking of diverse cargos by the SNX-BAR proteins. (To be submitted to *Nature Structural and Molecular Biology*)

Paul B, Hill T, Fairlie D, Collins BM. (2018) Designed cyclic peptides as inhibitors of the SNX-BAR proteins. (To be submitted to *Journal of Medicinal Chemistry*)

Paul B, Collins BM. (2018) Interaction and GAP activity of the SNX-RGS proteins with $G\alpha_s$. (To be submitted to *Biochemical Journal*)

Conference abstracts

Structural basis for the hijacking of endosomal sorting nexin proteins by *Chlamydia trachomatis*.
ASCB-EMBO Annual Meeting, 2017. Poster
East Coast Protein Meeting, 2017. Talk
Structural Biology and Biochemistry Theme meeting, 2017. Talk
COMBIO, 2016. Poster

Structural biology of a novel PX-protein that regulates GPCR signaling. East Coast Protein Meeting, 2015. Poster

Contributions by others to the thesis

The Chapter 2 presented in this thesis is a published research article. Author contributions are clearly stated in the corresponding sections of the respective chapter.

Assoc. Prof. Brett Collins, the primary supervisor of this work, provided the intellectual input for the initial research plan, directed and coordinated the development of all projects presented in this work, and edited and reviewed this thesis and all published material. As secondary supervisor, Assoc. Prof. Rohan Teasdale together with Dr. Marcus Kerr and Hyung Sung Kim contributed to the editing of manuscript and coordinated cell biology studies for the paper (Chapter 2). Dr. Timothy Hill and Prof. David Fairlie provided the synthetic cyclic peptides for the protein interaction studies (ITC) (Chapter 5), and Assoc. Prof Mike Henne and Ryan Feathers provided the bacterial expression constructs for the experiments in Chapter 7. Dr. Suzanne Norwood provided guidance in cloning baculovirus constructs used in my research work and performed the insect cell protein expression of these constructs. Dr. Thomas Clairfeuille provided technical guidance in ITC used in my research.

Statement of parts of the thesis submitted to qualify for the award of another degree

No works submitted towards another degree have been included in this thesis.

Research Involving Human or Animal Subjects

No animal or human subjects were involved in this research.

Acknowledgements

As I am getting ready to submit my Phd thesis, I am grateful to God for the amazing experiences I had, the wonderful people I met and for giving me the great opportunity to do research in the field that I desired.

I am forever thankful to my supervisor and my mentor Assoc. Prof. Brett Collins. I thankful to you for taking me under your wing and for providing your patient guidance and exceptional scientific inputs throughout my PhD. Maybe more than this, I am grateful to you for mentoring me and for proving to me that integrity, humility and honesty do exist in science. I hope and pray that I would remember these qualities wherever my career takes me and would be half as inspiring to others as you have been to me.

I want to thank all the Collins group members for being a bunch of enthusiastic, helpful and happy people to do science with. Thank you Dr. Saroja Weeratunga for all the scientific and non-scientific advice you showered me with. You were a great friend throughout and I am grateful to you.

I never had to worry about the administrative paperworks during my PhD due to our amazing postgraduate coordinators Dr Amanda Carozzi and Cody Mudgway. Thank you for all your help and support.

Papa, I want to thank you for your love, support and endless phone calls asking me whether I got any new crystals. This thesis is the result of your continuous encouragement and your constant prayers behind the scene.

I cannot thank enough my dearest husband and best friend, Sunil, who has been my rock. Thank you for moving all the way to Australia and putting your dreams on hold for a bit so I can have mine. Thank you for holding down the fort when I was busy finishing up my PhD, while growing a human inside me. I am grateful to you for everything and I love you.

Finally, let me thank my loving in-laws, my siblings, my friends and everyone who have been there to give me that little push and encouragement whenever I needed it, which led me to see through to the end of this PhD experience.

Financial support

This research was supported by the University of Queensland Research Scholarship (UQRS).

Keywords

sorting nexin, pathogen invasion, trafficking, transmembrane cargo, cyclic peptides, rgs domain, px domain, pxa domain, itc, x-ray crystallography

Australian and New Zealand Standard Research Classifications (ANZSRC)

ANZSRC code 060108 Protein Trafficking, 30%

ANZSRC code 060112 Structural Biology (incl. Macromolecular Modelling), 40%

ANZSRC code 060199 Biochemistry and Cell Biology not elsewhere classified, 30%

Fields of Research (FoR) Classification

FoR code 0601 Biochemistry and Cell Biology, 80%

FoR code 0699 Other Biological Sciences, 20%

Dedication

This thesis is dedicated to the memory of my mother, an extraordinary and humble woman whom I miss everyday.

Table of Contents

List of figures

Chapter 6

Chapter 7

Chapter 8

List of tables

List of Videos

List of abbreviations

5-HT6R	5-hydroxytryptamine subtype 6 receptor
AD	Alzheimer's disease
APP	Amyloid precursor protein
ARC	Apoptotic repressor with caspase recruitment domain
BAR	Bin-amphiphysin-rvs
cAMP	Cyclic adenosine monophosphate
CD	Circular dichroism
CIMPR	Cation independent mannose-6-phosphate receptor
DTT	Dithiothreitol
EB	Elementary body
ECL	Enhanced chemiluminescence
EDTA	Ethylenediaminetetraacetic acid
EEA1	Early endosome antigen 1
EGFR	Epidermal growth factor
ER	Endoplasmic reticulum
FPLC	Fast protein liquid chromatography
GAP	GTPase activating protein
GDP	Guanosine diphosphate
GFP	Green fluorescent protein
GLUT4	Glucose transporter 4
GPCR	G-protein coupled receptors
GPPNHP	5'-Guanylyl imidodiphosphate
GST	Glutathione-S-transferase
GTP	Guanosine triphosphate
HPLC	High-performance liquid chromatography
Incs	Inclusion membrane proteins
IGF1R	Insulin-like growth factor 1 receptor
IPTG	Isopropyl-β-D-thiogalactopyranoside
ITC	Isothermal titration calorimetry
K_d	Binding affinity
LAMP1	Lysosomal-associated membrane protein 1
LB	Luria Bertani
LD	Lipid droplet
LDL	Low-density lipoprotein

LDLR	Low-density lipoprotein receptor
MALLS	Multi-angle laser light scattering
Mdm1	Mitochondrial distribution and morphology 1
NMR	Nuclear magnetic resonance
NVJ	Nucleus-Vacuole Junction
OD	Optical density
PCC	Pearson correlation coefficient
PDB	Protein data bank
PI	Phosphoinositide
PI3K	Phosphatidylinositol-3 kinase
PM	Plasma membrane
PPI	Protein-protein interaction
PtdIns	Phosphatidylinositol
PtdIns3*P*	Phosphatidylinositol-3-phosphate
PtdIns(4,5)P_2	Phosphatidylinositol-4,5-bisphosphate
PX	Phox homology
PXA	Phox homology-associated A
PXC	Phox homology-associated C
RB	Reticulate body
RGS	Regulator of G-protein signaling
RMS	Root mean square
RTKs	Receptor tyrosine kinases
SDS-PAGE	Sodium dodecyl sulfate–polyacrylamide gel electrophoresis
SEC	Size-exclusion chromatography
SEMA4C	Semaphorin- 4C precursor
SNX	Sorting nexin
TB	Terrific Broth
TfR	Transferrin receptor
TGF-β	Transforming growth factor-β
TKR	Tyrosine kinase receptors
TM	Transmembrane
TGN	Trans-Golgi network
ULP1	Ubl-specific protease 1
VPS	Vacuolar protein sorting
WASH	Wiskott-Aldrich syndrome and SCAR homolog

Chapter 1

Chapter 1

Introduction

1.1 Endocytosis and the endosomal system

In 1963, Christian de Duve formulated the word “endocytosis” to describe the process by which eukaryotic cells internalise proteins, lipids, large particles and fluids from the cell surface into the cell cytoplasm (1). Endocytosis is essential for maintaining cellular homeostasis. Defects in endocytosis and underlying membrane trafficking pathways can lead to many different human diseases. The most well characterised endocytic process is receptor-mediated or clathrin-dependent endocytosis. However, the ultimate destination of *all* endocytic carriers is the intracellular endosomal system, where endocytosed material is sorted into different pathways of recycling, processing and degradation.

1.1.1. Receptor-mediated endocytosis and protein sorting

Receptor-mediated endocytosis or clathrin-dependent endocytosis provides a method for selective uptake of specific macromolecules, and this process involves the inward budding of certain domains of the plasma membrane. The macromolecules to be internalised are recognised by specific endocytic machinery, consisting of adaptor proteins that recognise structural motifs in the cytoplasmic domains of the cargo proteins that link them to the polymeric scaffolding protein called clathrin. This leads to formation of membrane invaginations called clathrin-coated pits, which bud away from the plasma membrane through the action of the dynamin GTPase protein. Cargo-containing vesicles move into the cell and fuse with the early endosomes (or sorting endosomes), an endomembrane compartment, which is enriched in the membrane lipid phosphatidylinositol-3-phosphate (PtdIns3*P*) (2,3).

In the early endosomal compartment, endosome-associated proteins sort the cargoes to be delivered to their appropriate destinations. Cargoes are either recycled back to the cell surface for reuse (directly from sorting endosomes or indirectly through recycling endosomal carriers), or to the trans- Golgi network (which allows access to the secretory pathway), or are degraded in lysosomes (where the cargoes are transported from early endosomes to the lysosomes through late endosomes). Thus early endosomes or sorting endosomes play an important role as a sorting station where the fates of internalised proteins and lipids (regardless of the mode of internalisation) are determined (3-5).

Many receptors and ligands that are internalised using receptor-mediated endocytosis have been studied in the context of endocytic recycling. Among the best characterised include proteins involved in nutrient uptake such as low-density lipoprotein receptor (LDLR) (**Fig. 1**), the transferrin receptor (TfR), glucose transporter 4 (GLUT4), and proteins involved in signalling such as G-protein couples receptors (GPCRs), and receptor tyrosine kinases (RTKs) (4,5).

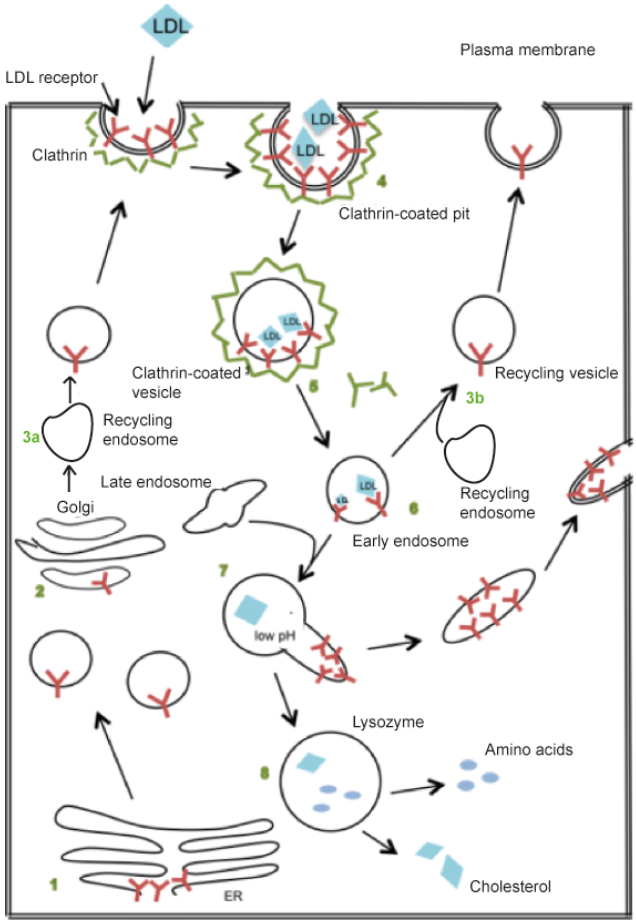


Figure 1*. Schematic diagram depicting the fate of a typical membrane protein low-density lipoprotein receptor (LDLR) 1) LDLRs are synthesised by ribosomes and folded in the endoplasmic reticulum (ER), 2-3) receptors are then modified in the Golgi and (3a) transported to the plasma membrane enroute recycling 4) when the LDL ligand binds to LDLR on the cell membrane the receptor- ligand complex gets internalised in a clathrin-coated pit, 5) this pinches off and becomes a clathrin-coated vesicle which then moves into the cell and 6) enters the early endosome and some are directly moved to the recycling endosomes (3b) which gets recycled back to the plasma membrane directly through recycling endosome or 7) this endosome fuses with the late endosome and because of the low pH (around 5) LDL particles dissociate from LDLRs. A vesicle or tubule that contains recycling receptors buds off and returns the LDLR back to the plasma membrane, 8) the late endosome containing the LDL fuses with the lysosome where LDL is degraded to release cholesterol and amino acids (scheme adapted from (6)).*

1.2 Significance of endosomal trafficking

Endocytic recycling and trafficking pathways control the balance between the late endosomal degradation and cargo retrieval, which in turn maintains the proper protein and lipid composition of the plasma membrane and various organelles such as lysosomes and the Golgi. Thus it is crucial in many major cellular processes including cell signaling, nutrient uptake, adhesion and all forms of cellular homeostasis (4). While our understanding of endosomal sorting and signaling has improved in recent years, it still remains limited and many questions remain unanswered.

The importance of endosomal homeostasis is highlighted by the fact that many different human diseases are caused by defects in endosomal trafficking and regulation (7). Alterations in expression or mutations in a number of proteins involved in endosomal sorting are responsible for a variety of disorders related to dysfunction in specific endosomal pathways. One of the major protein families that play a critical role in endosomal sorting and cell signaling are the sorting nexins (SNXs) (7,8)

1.3 Sorting nexins (SNXs) and endosomal sorting

SNX proteins are members of a larger family of proteins called PX-domain proteins, defined by the presence of a phosphatidylinositol phospholipid (or phosphoinositide)-binding module; the Phox-homology (PX) domain.

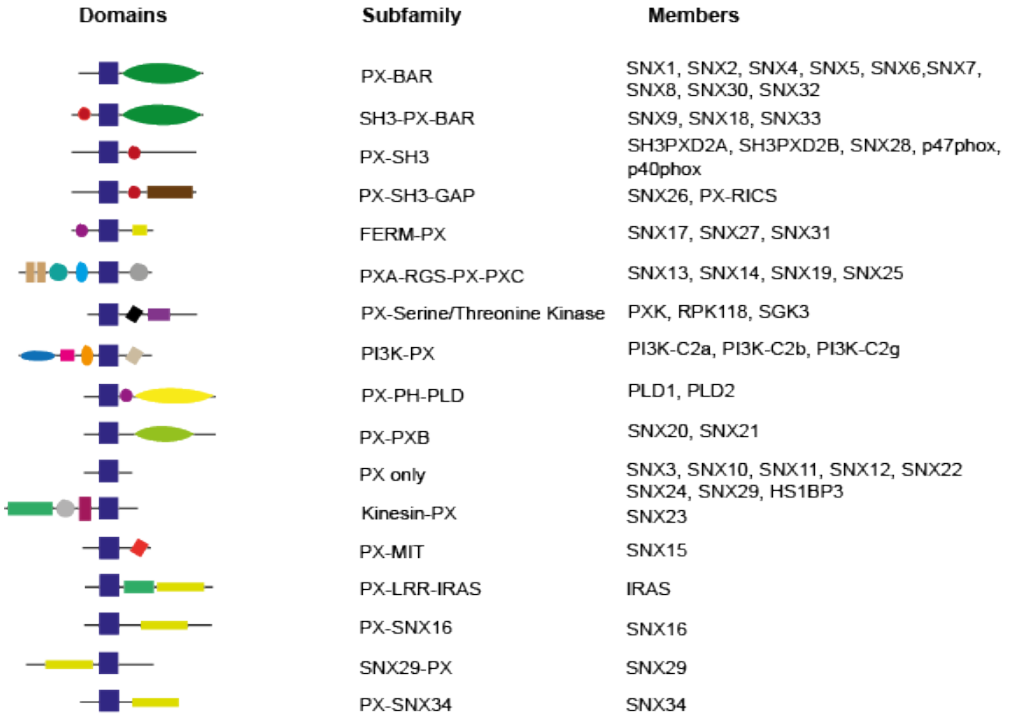


Figure 2. *Classification of PX-domain proteins. Adapted from Reference (8).*

Currently there are 49 PX domain-containing proteins identified in the mammalian genome. Teasdale and Collins (2012), published a revised classification of PX-domain proteins based on sequence homology and known domains, as well as their conserved secondary structure predictions (8) (**Fig. 2**). The best-characterised role of the PX domain is to bind to phosphoinositide lipids and promote the association of the PX-domain proteins with various membranes of the endocytic and secretory system. The most common lipid bound by the PX domain is phosphatidylinositol-3-phosphate (PtdIns3*P*). Thus this domain is frequently involved in targeting SNXs to early endosomal compartments that are enriched in PtdIns3*P* due to the activity of PtdIns-3 lipid kinases. SNXs play critical roles in endocytosis, endosomal sorting and membrane trafficking (**Fig. 3**).

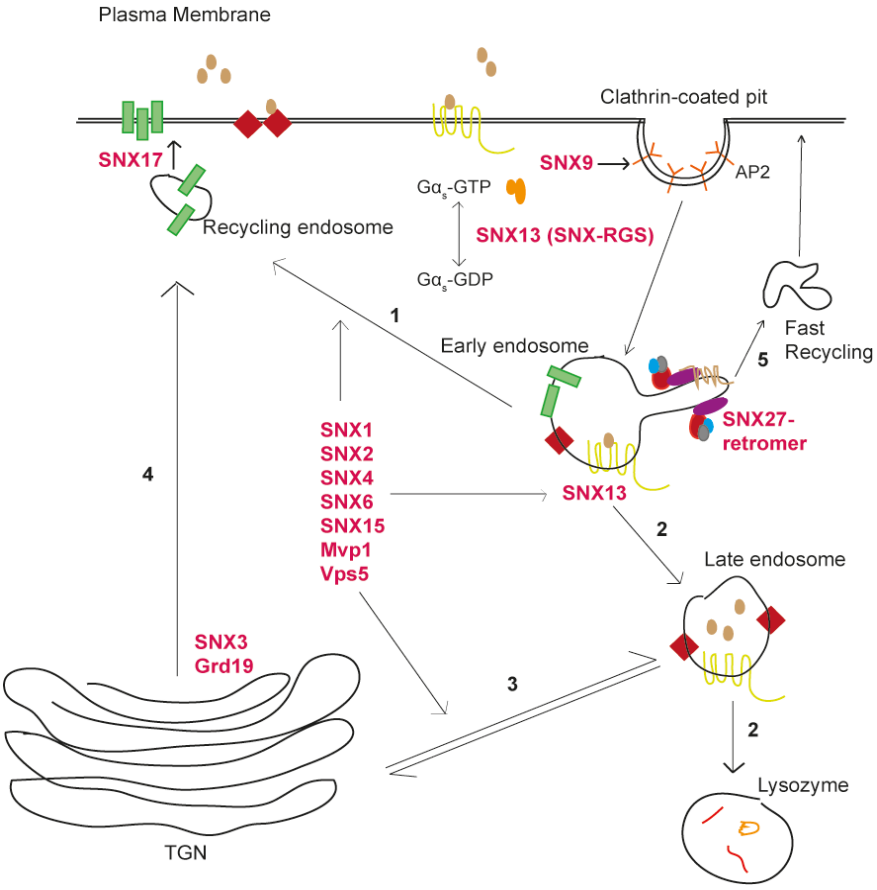


***Figure 3**. Some of the known roles of sorting nexins (SNXs) in endosomal sorting and membrane signalling. 1) To its return to the plasma membrane, Transferrin receptor (TfR) is sorted into recycling endosomes. 2) Through path 1, certain tyrosine kinase receptors (TKRs) and G-protein-coupled receptors (GPCRs) are recycled. Sometimes they are seen in late endosomes (2) to be transported for lysosomal degradation. 3) Pathway of hydrolases to and from the trans-Golgi network (TGN) to the late endosome, which are transported to the lysosome. 4) Sorting of proteins in the TGN for delivery to the plasma membrane. 5) SNX27-retromer promotes a fast recycling route where it transports cargoes containing PDZbms from endosomes to the plasma membrane.*

Most SNXs contain one or more additional domains other than the PX domain (**Fig. 2**), which allows them to perform a wide variety of functions that include mediating protein-protein interactions, membrane remodelling, and lipid metabolism, as well as many other functions that remain unknown (7,8). With the fundamental roles of SNXs in endosomal sorting and signalling, comes their link to many diseases such as cancer, neurodegeneration, heart disease and chromosomal disorders (5,7,8).

During my PhD I have focussed on studying the structure and function of SNX proteins from two distinct subfamilies.

(ii) The first of these are the SNX-BAR proteins SNX5, SNX6 and SNX32, which contain a PX domain and an additional Bin/amphiphysin/Rvs (BAR) domain. Recent evidence suggests that these proteins might act as specific adaptors that bind and transport selected transmembrane cargos. Effector proteins that are secreted by the bacterial pathogen *Chlamydia trachomatis* also hijack specific SNX-BAR proteins.

(i) The second subfamily I have studied are the unique transmembrane SNX-RGS proteins, SNX13, SNX14, SNX19 and SNX25, which play an important role in mediating interactions between the membranes of the ER, lipid droplets (LDs), and endosomal compartments, and potentially act as GAP proteins for $G\alpha_s$ subunits.

Further details about these SNX sub-families are given in the following sections.

1.4 The SNX-BAR protein family

The largest and arguably best-studied sub-family of SNX proteins is the SNX-BAR sub-family. This group consists of SNX1, SNX2, SNX4, SNX5, SNX6, SNX7, SNX8, SNX30 and SNX32 (**Fig. 2**). The N-terminal PX domain within these proteins is generally believed to be the primary phosphoinositide-binding domain directing specific organelle recruitment. The C-terminal BAR (Bin/Amphiphysin/Rvs) domain is a protein motif that consists of three α-helices that dimerises and forms a rigid banana-shaped coiled-coil structure (**Fig. 4**). The concave surface is composed of positively charged basic residues that associate with negatively charged membrane phospholipids through electrostatic interactions. When coupled to the phosphoinositide-binding PX domain, the dimeric BAR domain structure can promote membrane curvature leading to membrane tubulation. This sub-family is evolutionarily conserved from yeast to humans and is important in endocytosis as well as in membrane sorting events (8-10). It is believed that different combinations of

homodimeric and heterodimeric SNX-BAR proteins may control different sorting events within the cell, although this remains poorly understood.

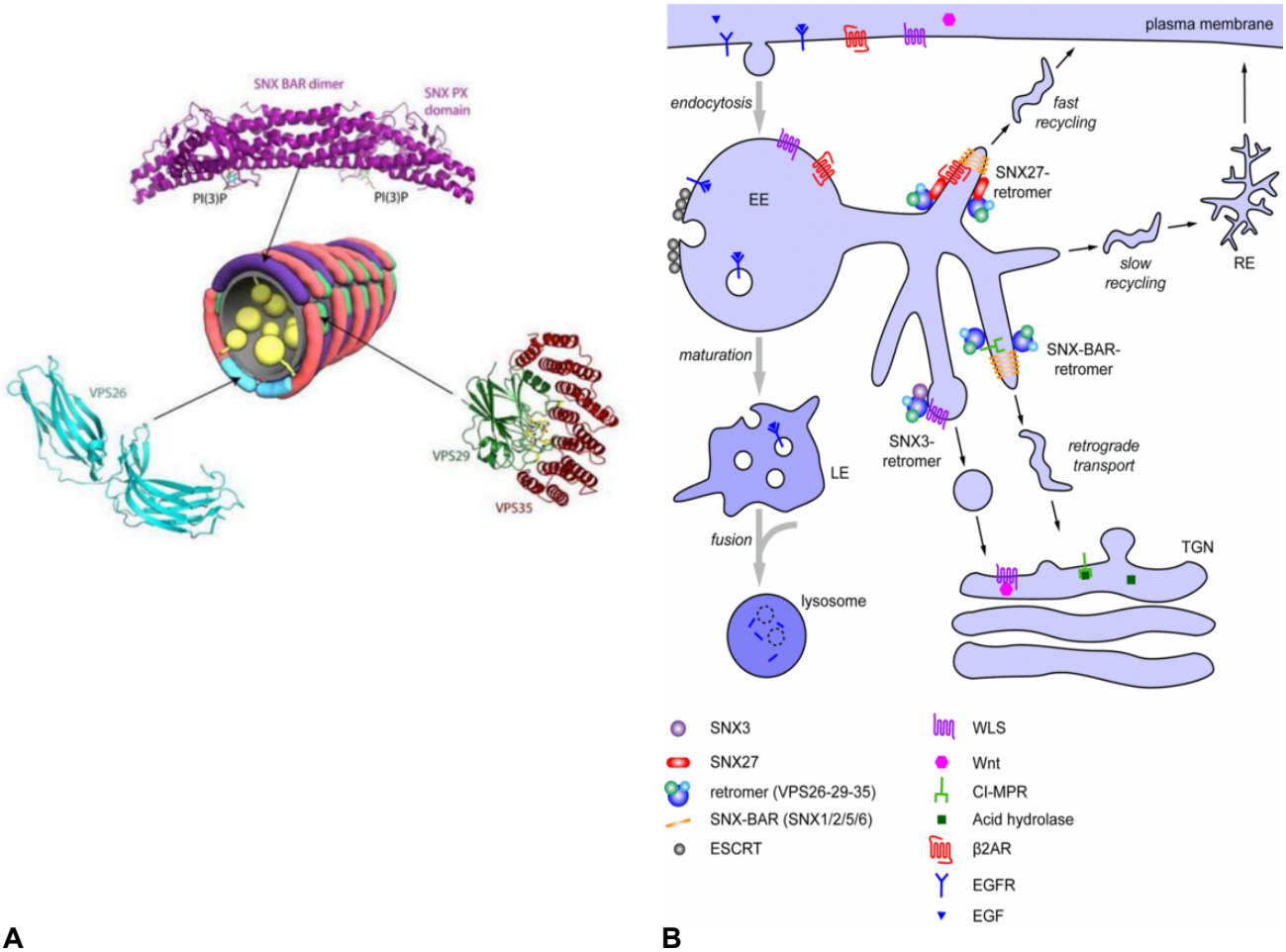
One of the major functions of the SNX-BAR proteins is to regulate tubular-based endosomal sorting. SNX1, SNX2, SNX5 and SNX6 form specific heterodimeric combinations, where SNX1 or SNX2 form heterodimers with SNX5 or SNX6 (11). A third brain specific homologue of SNX5 and SNX6 called SNX32 may function similarly, but it has not been studied to date. Their role in endosomal trafficking is often coupled to their interaction with a trimeric protein complex called 'retromer'. Early studies on retromer demonstrated their role in regulation of cargo proteins from endosomes to the *trans*-Golgi network (TGN), allowing cargo retrieval away from late endosomal/lysosomal degradative pathway (12,13). Out of the many retromer cargoes, the best studied are the acid hydrolase receptors, cation-independent mannose-6–phosphate (CI-MPR) in mammals, and vacuolar protein sorting 10 (Vps10) in yeast. These receptor cargos are essential for lysosomal activity because they deliver acid hydrolase enzymes from the TGN to the vacuole (in yeast) or lysosome (in mammals). After releasing their substrates, these receptors recycle back to TGN via retromer/SNX-BAR interactions to facilitate future cargo-hydrolase transportation (14). Retromer/SNX-BAR complexes are also important for transport of other cargoes such as signalling molecules like Wntless (receptor for Wnt morphogens), SNAREs and SorLA (receptor for amyloid precursor protein) (9,10).

1.4.1 The Retromer Complex

A major regulator of retrograde trafficking (endosome-to-Golgi retrograde retrieval) is the SNX-BAR/retromer complex. In yeast this consists of the retromer subunits Vps26p, Vps29p, Vps35p, coupled to the heterodimeric SNX-BAR complex of Vps5p and Vps17p. Unpublished evidence from our lab shows that Vps26p, Vps29p and Vps35p (trimer) form an exterior scaffolding subcomplex, while Vps5p and Vps17p are the yeast SNX-BAR homologues that dimerise and mediate membrane interaction and remodeling (B. Collins, unpublished), which is slightly different to the schematic model shown in **Fig. 4**.

In mammals, the SNX-BAR dimer binds to specific phosphatidylinositol lipids to determine its subcellular membrane association and thus helps in recruiting the Vps trimer to endosomal compartments. The mammalian counterpart of retromer consists of the Vps trimer (Vps26, Vps29 and Vps35) and alternative membrane interacting SNX-BAR dimers. These dimers consists of the Vps5p orthologues SNX1 or SNX2 and the Vps17p orthologues SNX5 or SNX6 (10,17). In the SNX dimer, SNX1 and SNX2 are interchangeable Vps5 orthologs (17). The highly similar proteins, SNX5 and SNX6, are

shown to have overlapping functions where they both have similar roles in the retrograde sorting of mannose-6–phosphate receptor, and they both form heteromeric complexes with the retromer-associated SNX1 protein. Thus SNX5 and SNX6 are thought to be functionally interchangeable orthologs of Vps17 in mammalian cells (15,18).



A **B**

Figure 4. *Current model of SNX-BAR/retromer-mediated trafficking.* ***A****) In mammals SNX-BAR heterodimers (SNX1/SNX5 or SNX1/SNX6; SNX2/SNX5 or SNX2/SNX6) form transport coats in complex with the cargo-selective VPS26-VPS29-VPS35 subcomplex. SNX-BAR proteins promote the endosomal membrane curvature resulting in the formation of membrane tubules that transport cargo to the trans-Golgi network (TGN) (adapted from (15));* (***B***) *Based on the accessory protein bound, there are three distinct retromer mediated trafficking pathways. One of the examples shown here is the SNX-BAR-retromer mediated retrograde trafficking of cargo called CI-MPR from the endosomes to the TGN (16).*

The retrograde machinery is important to maintain the distribution of lipids and proteins in endosomes and the TGN. Defects in retromer function are known to be involved in neurodegenerative disorders including Alzheimer's and Parkinson's (19-22).

Although currently it is not entirely clear which specific neuronal trafficking pathways are defective in these patients, the build-up of toxic species due to perturbed lysosomal function is believed to play an important part in disease progression. In addition, several plant, bacterial and viral proteins are known to depend on endocytosis and retromer-dependent trafficking for their host cell entry and delivery to target organelles (10).

1.4.2 The unusual structure of the SNX5 PX domain

The PX domain is a phosphoinositide-binding domain, approximately 120 amino acids in length, and composed of a three-stranded β-sheet followed immediately by three α-helices. The first and second α-helices are linked by an extended structure that is rich in Proline residues, called the polyPro loop (ΨPxxPxK loop; Ψ is a hydrophobic side-chain). Previous crystal structures of PX domains of SNX9 (23) and p40phox (24) in complex with PtdIns3*P* reveal the presence of four side chains that are required for binding the headgroup of the endosomal phosphoinositide lipid PtdIns3*P*. They are Arg, Tyr, Lys and Arg residues marked in **Fig. 5** as 1, 2, 3 and 4 respectively. These residues are conserved in most PX domains including in SNX1 and SNX2. However, they are completely absent in SNX5, SNX6 and its closely related homologue SNX32. Moreover, SNX5, SNX6 and SNX32 contain a 30-residue insertion immediately after the ΨPxxPxK loop region that structurally forms an extended helix-turn-helix motif (**Fig. 5** and **Fig. 6**) (8,25). These features set these proteins apart from the rest of the SNX family members.

Figure 5. *Amino acid sequence alignment of PX- domains of mammalian SNX- families.*

The X-ray crystal structure of the SNX5-PX domain confirms the lack of conserved residues required to form a pocket for PtdIns3*P*-binding, and in addition shows the structural insertion of the extended α-helical region obstructs the typical phospholipid

interaction site. Together these features suggest an inability to interact with phosphoinositide lipids in the same way that other PX-domain proteins do. Although there is some evidence from nuclear magnetic resonance (NMR) spectroscopy experiments for the weak association of the SNX5 PX domain with the lipid PtdIns(4,5)P_2 (25) the structure of the SNX5 PX domain suggests this is likely an artefact and not functionally significant (**Fig. 6B**) It has remained unclear what the functional significance of this α-helical insertion is, other than it may contribute to an altered membrane binding and remodeling activity by these SNX-BAR proteins when forming heterodimers with SNX1 or SNX2. As described in this thesis however, I have now shown that it plays a critical role as a protein interaction site.

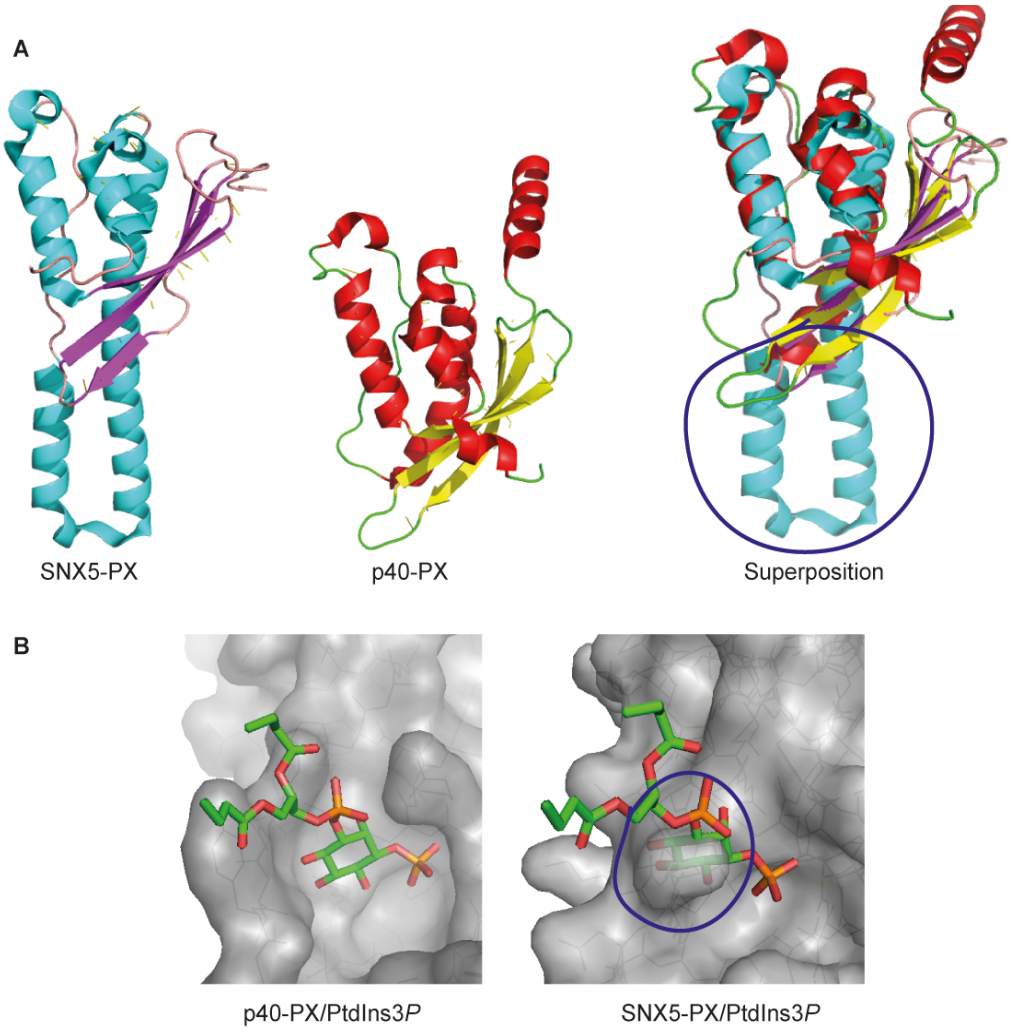


***Figure 6**. **(A)** X-ray crystal structures of SNX5PX domain (PDB: 3HPB) and p40-PX domain (PDB: 1H6H) and their best fit superposition (on the right) showing the most striking difference of α-helical insertion region (circled) in SNX5-PX domain; **(B)** Phosphoinositide binding sites in the (left) p40-PX/PtdIns3P complex that shows a perfect PtdIns3P binding pocket and (right) a lack of binding pocket in SNX5-PX due to the obstruction caused by the α-helical region (figure idea taken from (25)).*

1.4.3 Chlamydia trachomatis, membrane inclusion proteins (Incs) and host cell trafficking

Although the SNX-BAR proteins play important roles in normal cellular function, two recent studies also indicate a potentially important role for SNX5 in host cell invasion by the bacterial pathogen *Chlamydia trachomatis* (26,27). As mentioned above, there are many plant, bacterial, and viral proteins that depend on endocytosis for their successful host cell entry and thus survival within the cell. Of the many pathogens that hijack retrograde trafficking pathways for effective infection *C. trachomatis* is one of the most striking examples, as well as being an important human pathogen that causes 100 million new infections each year worldwide. *C. trachomatis* is a major reason behind various human diseases for which no effective vaccine exists. The diseases include noncongenital blindness worldwide, sexually transmitted diseases and noncongenital infertility in Western countries (28,29).

Once inside the human cell, *C. trachomatis* replicates within a membrane-bound vacuole called the inclusion. The intracellular cycle of development is common among all Chlamydiae, and alternates between two bacterial forms; the elementary body (EB), which is infectious but non-dividing, and the reticulate body (RB), which is non-infectious but dividing. Upon internalization of EBs by endocytic pathways, the bacteria reside in the inclusion, which is separated from the canonical endo-lysosomal organelles preventing normal cellular degradation and shielding the bacteria from innate cellular immune factors. The EBs then differentiate into RBs and thus the replication commences inside the growing inclusion. After replicating for 24-72 hours inside the ever-enlarging inclusion, the RBs re-differentiate to EBs, which are then released into the neighboring cells ready to infect (27,30). The surrounding inclusion membrane acts as the interface between the host cell and the bacteria.

Chlamydia genomes, despite their small size, encode a large number of genes that encode secreted virulence effectors. Soon after the beginning of chlamydial protein synthesis, the inclusion membrane is extensively modified through insertion of numerous type III secreted effector proteins called inclusion membrane proteins (Incs). Incs are a large family of proteins, which are highly diverse and expressed only during the context of infection and are specifically localised to the inclusion membrane. Incs possess little primary amino acid sequence similarity, however they are all composed of a cytoplasmic N terminal domain and a cytoplasmic C-terminal domain, separated by two closely spaced transmembrane regions connected by a short loop (**Fig. 7)**. After inserting into the inclusion membrane, they position at the host-pathogen interface and extend their termini

into the host cytoplasm. There are 59 putative Incs in *C. trachomatis* and their functions are largely unknown. One function may be to interact with the host cell machinery to help with the chlamydial survival in the cell. Even though Incs are critical for the Chlamydial infection and the host-pathogen interactions, only a few host targets have been identified (31,32). As mentioned above it was recently found that SNX5 is targeted to chlamydial inclusions through the binding of one such inclusion protein called IncE (26).

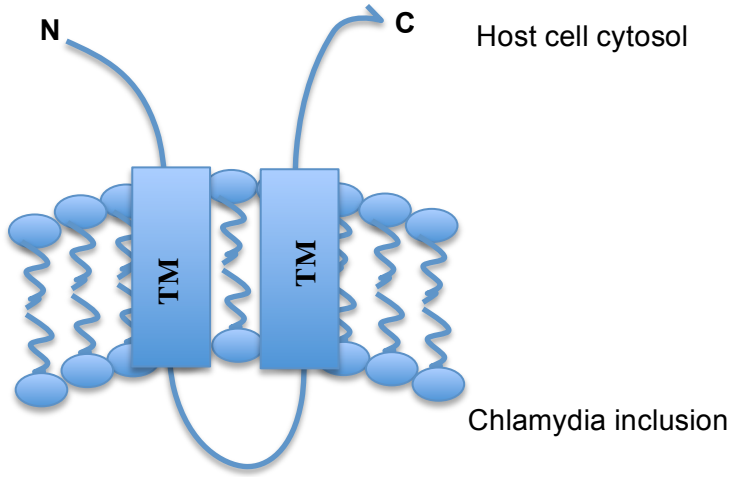


Figure 7. *Predicted topology of Inc proteins in Chlamydia inclusion membrane.*

1.4.4 SNX5 and SNX6 interact with the Chlamydial inclusion protein IncE

Recently Mirrashidi *et al*. reported a protein-protein interaction network between human host cell proteins and chlamydial Inc proteins identified by affinity purification and mass spectroscopy. This network revealed interactions between IncE (also called CT116) and SNX-BAR proteins that were subsequently confirmed *in vivo* and *in vitro*. These studies found that the IncE C-terminal cytoplasmic region (IncE 101-132) was necessary and sufficient to bind specifically to SNX-BAR proteins in immunoprecipitations. PSIPRED analysis suggested the 101-132 residues of IncE contain a β-hairpin structure, and the truncations demonstrated that this structure was required for SNX5 binding. IncE 101-132 demonstrated no binding to the PX-domains of SNX1 and SNX2, whereas it bound directly to the PX domains of SNX5 and SNX6 (26).

As described in the previous section SNX5 and SNX6 PX domains contain a unique 30 amino acid long α-helical insertion, which is absent in SNX1 and SNX2. We speculated that it is because of this distinction that IncE specifically recognises SNX5 and SNX6, and not SNX1 and SNX2 (26). Further investigation is required to understand the molecular details of this interaction.

The role of SNX5 and/or SNX6 in *C. trachamotis* infection is still not clear. IncE directly recruits SNX5 and SNX6 to the inclusion membrane to form membrane tubules (and SNX1 and SNX2 via indirect heterodimerisation), and it was found by Mirrashidi *et al*. (2015) that SNX5/SNX6 depletion increases infectious progeny production. This suggests a role for the SNX-BAR proteins in restricting chlamydial infection. Retrograde trafficking might be involved in the recognition or clearance of Chlamydia. Retromer is implicated in maintaining the integrity of the trans-Golgi. Thus the IncE-mediated recruitment of retromer SNX-BARs away from the endosomal compartments may play a role in Golgi fragmentation observed during Chlamydia infection (33). This process promotes lipid acquisition by *C. trachomatis*, which may lead to enhanced progeny production that is exacerbated upon further SNX5/SNX6 depletion (26,33). Mirrashidi *et al.* reported that IncE and SNX-BARs colocalise on the inclusion membrane and tubules, and that ectopic expression of IncE is sufficient to induce SNX-BAR-mediated membrane tubulation. Depletion of SNX-BAR components limits inclusion tubulation (34).

Studies on other intracellular pathogens have demonstrated that other strategies can also be employed to repurpose retromer and SNX-mediated trafficking. Silencing of retromer components abrogates HIV, HPV, Coxiella and Salmonella infections (26,35-37). In addition to retromer SNX-BAR proteins, Salmonella is found to recruit the VPS complex. On the other hand, Chlamydia IncE recruits only retromer SNX-BAR proteins. Based on this, Mirrashidi *et al*. developed a model of where IncE directly interacts with SNX5 and SNX6 to recruit the SNX-BAR proteins to the inclusion membrane, which can result in two functional consequences; a) inclusion membrane tubulation can be induced by the SNX-BAR proteins due to the functional ability of the proteins with BAR domains to induce membrane curvature, b) the segregation and recruitment of SNX-BAR proteins may result in the dysfunction of normal retromer functions, thus disabling them from imposing a restriction on *C. trachomatis* and thus the infection (**Fig. 8**).

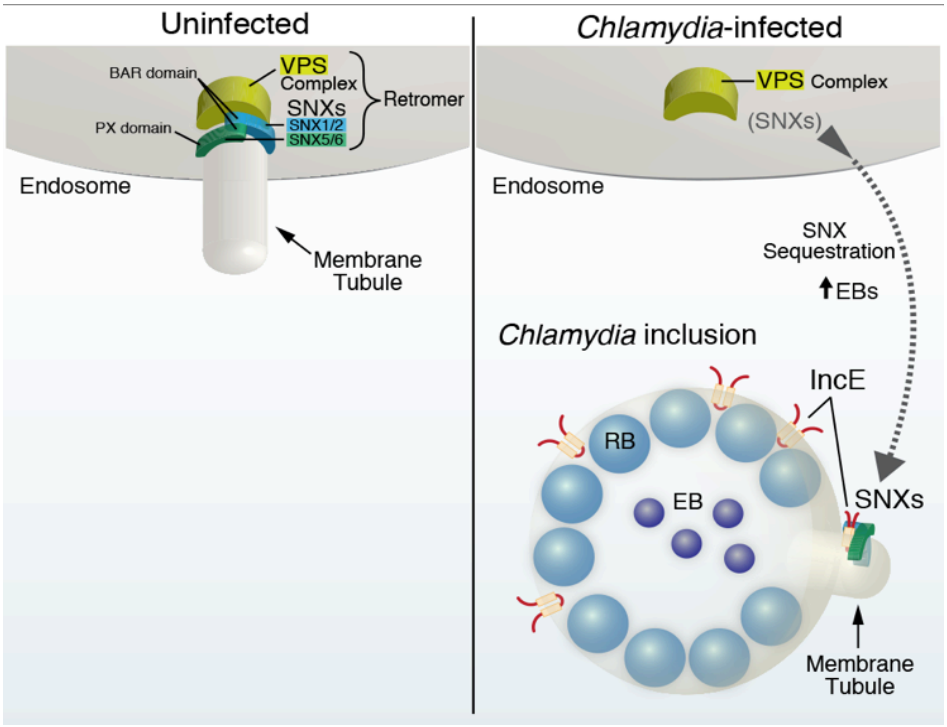


Figure 8. *Figure demonstrating the binding model of IncE to SNX5/SNX6 and possibly SNX32. From reference (26).*

1.5. The SNX-RGS protein family

The SNX-RGS proteins are a SNX sub-family that are conserved in all eukaryotes. The human genome encodes four homologues SNX13, SNX14, SNX19 and SNX25, while *Saccharomyces cerevisiae* possess a single homologue called Mdm1p. These proteins have a conserved architecture, composed of two N-terminal transmembrane helices, an N-terminal PX-associated domain (PXA), a regulator of G-protein coupled receptor (RGS) domain (present in all proteins except SNX19), a central PX domain, and a C-terminal PX-associated domain (PXC) (8,38) (**Fig. 9**). Note that the PXA and PXC domains are not related in sequence or structure to each other. Out of the four globular domains in SNX-RGS proteins, the PX and RGS domains are the best characterised, while the structures and the functions of the PXA and PXC domains are completely unknown.

Mitochondrial distribution and morphology 1 protein (Mdm1p) is the SNX-RGS protein found in *S. cerevesiae*. Secondary structure predictions of Mdm1p indicate the presence of identical domain organisation to mammalian SNX-RGS proteins. The *Drosophila melanogaster* snazarus protein also contains all the putative domains, thus making Mdm1p and snazarus evolutionary homologs of the mammalian SNX-RGS proteins (38).

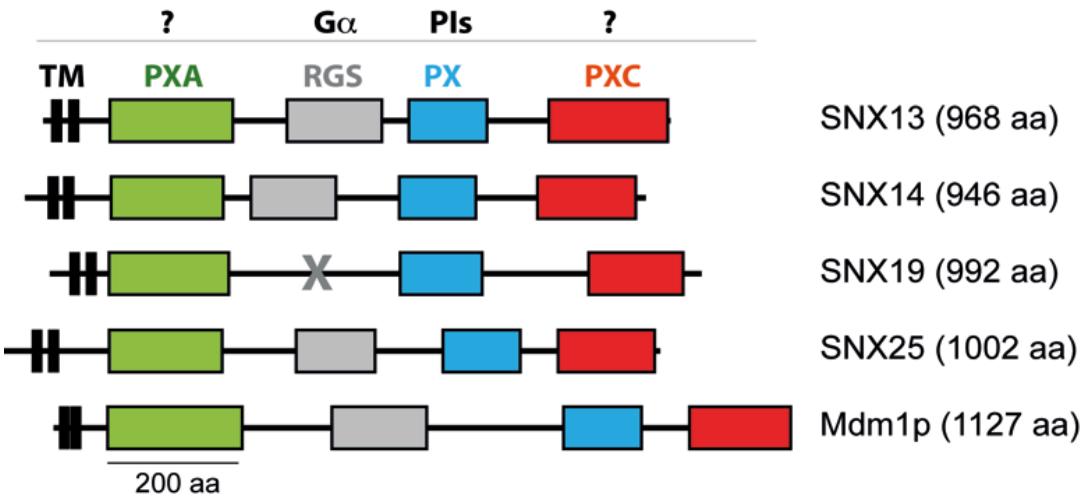


Figure 9. *Diagram of human (SNX13, SNX14, SNX19, SNX25) and yeast (Mdm1p) SNX-RGS proteins showing PXA, RGS, PX and PXC domains. Potential ligands for the domains are indicated above. Adapted from (38).*

1.5.1 Regulator of G-protein signalling (RGS) domain

G-protein coupled receptors (GPCRs) consist of seven transmembrane α-helical domains with an N-terminal extracellular and a C-terminal intracellular tail. Upon binding of a ligand the GPCRs are activated, which leads to a conformational change that allows the activation of associated Gα-GDP/ Gβγ complexes by the release of GDP and thus allowing the binding of GTP on the Gα subunit (39,40). When bound to GDP, the heterotrimeric GTP-binding proteins (G-proteins) are considered to be in the off state and when bound to GTP they are in the on or activated state that is able to bind to subsequent effector molecules. The GTP-bound state of Gα subunits controls the duration of G-protein signals downstream of the activation of associated GPCRs (**Fig. 10**). The RGS domain is a small, α-helical structure that serves as a GTPase-activating protein (GAP) for Gα subunits of heterotrimeric G-proteins. RGS domains bind to these activated Gα subunits and act as GAPs, which accelerates the hydrolysis of GTP to GDP and switches off GPCR signalling (**Fig. 10**) (38,41,42).

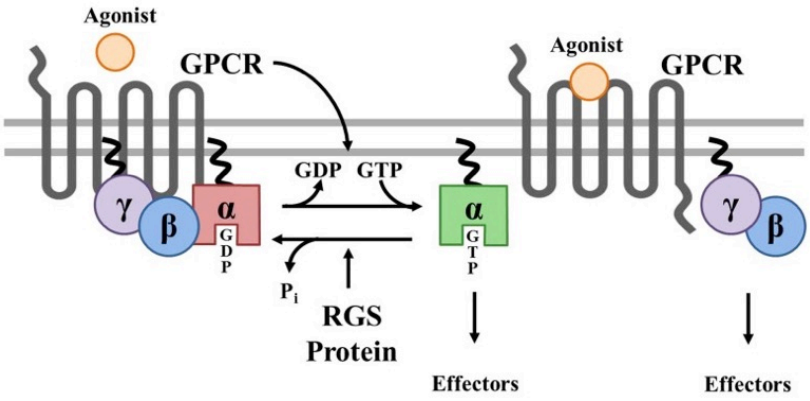


Figure 10. *Schematic diagram depicting how RGS proteins regulate GPCR signaling. Adapted from (43).*

1.5.1.1 Structural insight into the mechanism of intrinsic GTP hydrolysis by Gα subunits

The inactive Gα subunit interacts with GDP within a nucleotide-binding pocket that contains residues from both the Ras-like domain and the α-helical domain, which contains a six-helix bundle. When activated, the GDP gets exchanged for GTP and the residues in the binding pocket interact with the γ-phosphoryl group of the GTP, which induces a structural rearrangement within three switch regions (SI–SIII) of Gα. X-ray crystallographic data from Gα transition state-mimetic forms (i.e., Gα bound to GDP and AlF_4^-) have shed light into the structural mechanism behind the intrinsic GTP hydrolysis by Gα (44) (**Fig. 11**).

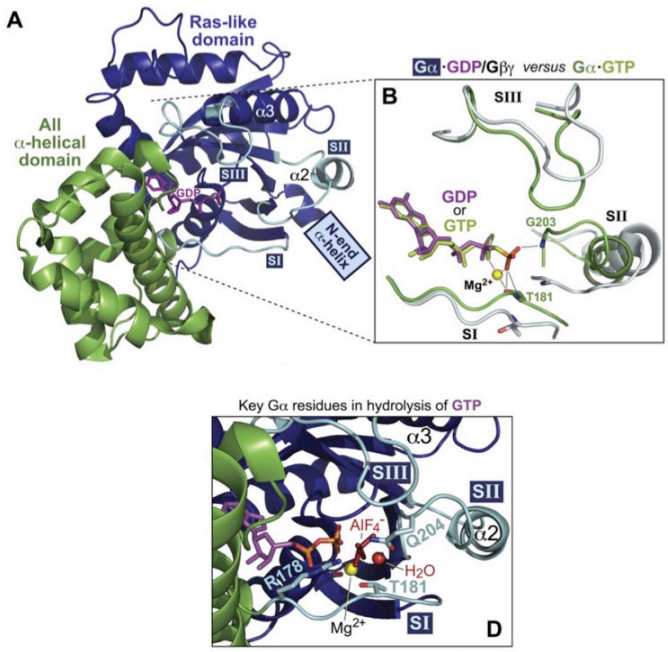


Figure 11. *Structural mechanism of intrinsic GTP hydrolysis by Gα subunits.* ***(A)*** *The GDP (magenta) binding site is located between the Ras-domain (blue) and the α-helical domain (green) of the Gα subunit (PDB: 1GP2).* ***(B)*** *Structural differences between GDP- and GTP-bound states.* ***(C)*** *Structural mechanism of GTP hydrolysis by Gα. Figure adapted from (45)*

The hydrolysis of GTP reaction is promoted by three Gα amino acids that are conserved. The critical nucleophilic water molecule required for the GTP γ-phosphate hydrolysis is coordinated by the Gln-204 residue in switch II, whereas the switch I residues Arg-178 and Thr-181 stabilizes the leaving group and the Thr-181 coordinates a bound Mg^{2+} ion (44) (**Fig. 11**).

1.5.1.2 Structural insights into how RGS domain stabilise a Gα subunit

The RGS domain is composed of nine α-helices that bind to the Gα transition state for GTP hydrolysis. The nine helices consists of two subdomains, the first one is composed of helices α1, 2, 3, 8, and 9; the second subdomain consists of helices α4, 5, 6, and 7. Each subdomain is antiparallel to each other. The unrelated GAPs for small G-proteins have a critical arginine at the active site that contributes to the GTP hydrolysis. In contrast, RGS proteins do not contribute a specific residue to the nucleotide-binding pocket for the catalytic mechanism. X-ray crystallographic and NMR studies have shed light into the structures of RGS proteins, as well as the mechanism by which they interact with and regulate Gα proteins. Gα-GDP bound with the planar ion aluminum tetrafluoride (AlF_4^-) mimics the transition state of Gα (GTP → GDP + P_i). RGS proteins are typically selective for binding to this transition state of Gα (46,47).

There are three critical contacts formed between RGS proteins and Gα. As shown in Fig. 12, the side chain of Asn-122 residue in RGS8 forms a hydrogen bond with the important Gln-204 residue of Gα, which is responsible for GTP hydrolysis. This hydrogen bond enables the Gln-204 residue to stabilize the GTP γ-phosphate that is being hydrolyzed. Asn-82 residue (RGS8) interacts with the hydroxyl group of Thr-182 ($Gα_{i3}$) residue from switch I thus enabling the side-chain hydroxyl to contact the Lys-210 ($Gα_{i3}$) residue from switch II. These interactions locks the Gα switches I and II into their transition state conformations, thus allowing accelerated GTPase activity. An Asp-157 residue (RGS8) that is conserved in all RGS proteins (except RGS2), functions as the stabilizer of the backbone amine of the Gα Thr-182 residue (switch I) to allow the side-chain hydroxyl group of the adjacent Thr-181 residue to stabilize the Mg^{2+} cation [35](47) (**Fig. 12**).

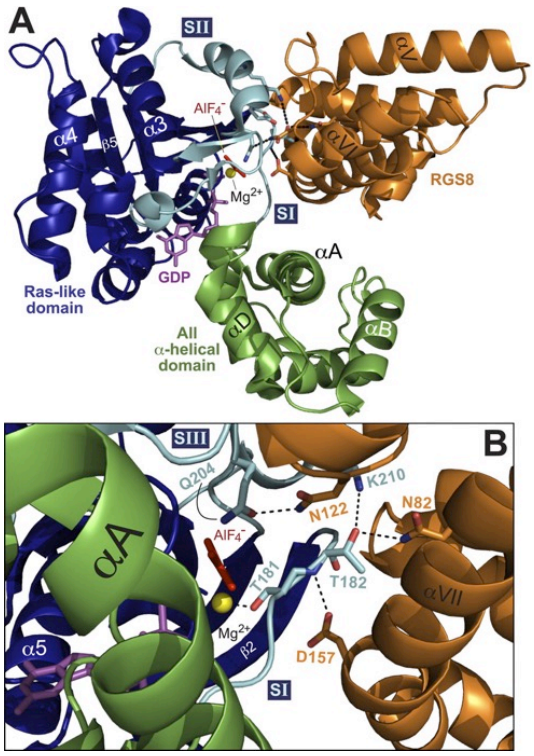


Figure 12. *Structural mechanism behind the stabilisation of Gα by RGS proteins using a diagram of the RGS8/Gα_{i3} structure. The Gα_{i3} α-helical domain (green), the Ras-like domain (dark blue), the three switch regions (SI, SII, and SIII) (cyan), the nucleotide (magenta), AlF_4^- (red), Mg^{2+} (yellow) and RGS8 (orange) are shown in the figure. Figure adapted from (45).*

1.5.1.3. The RGS domains of SNX-RGS proteins

All proteins except SNX19 have the RGS domain upstream of the PX domain (**Fig. 9**). Previous studies have reported that SNX13 has GAP activity specific for the Gα_s isoforms of the heterotrimeric G-proteins, which are involved in cAMP signalling from GPCRs such as serotonin and β-adrenergic receptors (48). SNX13 is shown to bind Gα_s and promote a 20-fold increase in the intrinsic catalytic rate of GTP hydrolysis; which in turn attenuates Gα_s-mediated signalling (48). Until recently it was unknown whether the RGS domain from other SNX-RGS proteins share the same GAP activity as SNX13. Ha *et al.* (2015) studied the activity of the SNX14 RGS domain and found that it does not appear to share this GAP activity. But it did appear to bind specifically to and sequester Gα_s and thus inhibit downstream cAMP production, essentially resulting in a similar functional outcome to activating GTP hydrolysis (49). Whether the SNX25 RGS domain possesses GAP or Gα binding activity remains unknown, but has been reported to be involved in transforming growth factor-β (TGF-β) signalling, lysosomal trafficking and degradation of TGF-β receptor 1 (TβRI) by regulating interaction of with TβRI (38,49).

1.5.2 Structural insight into the PX domains of SNX-RGS proteins

Even though most PX domains are commonly thought to bind the early endosomal lipid marker PtdIns3*P*, our lab showed that mammalian SNX-RGS proteins have a wider range of phosphoinositide binding preferences. This may direct recruitment to specific membranes, to control their roles in membrane trafficking and signalling at different cellular compartments (8,38).

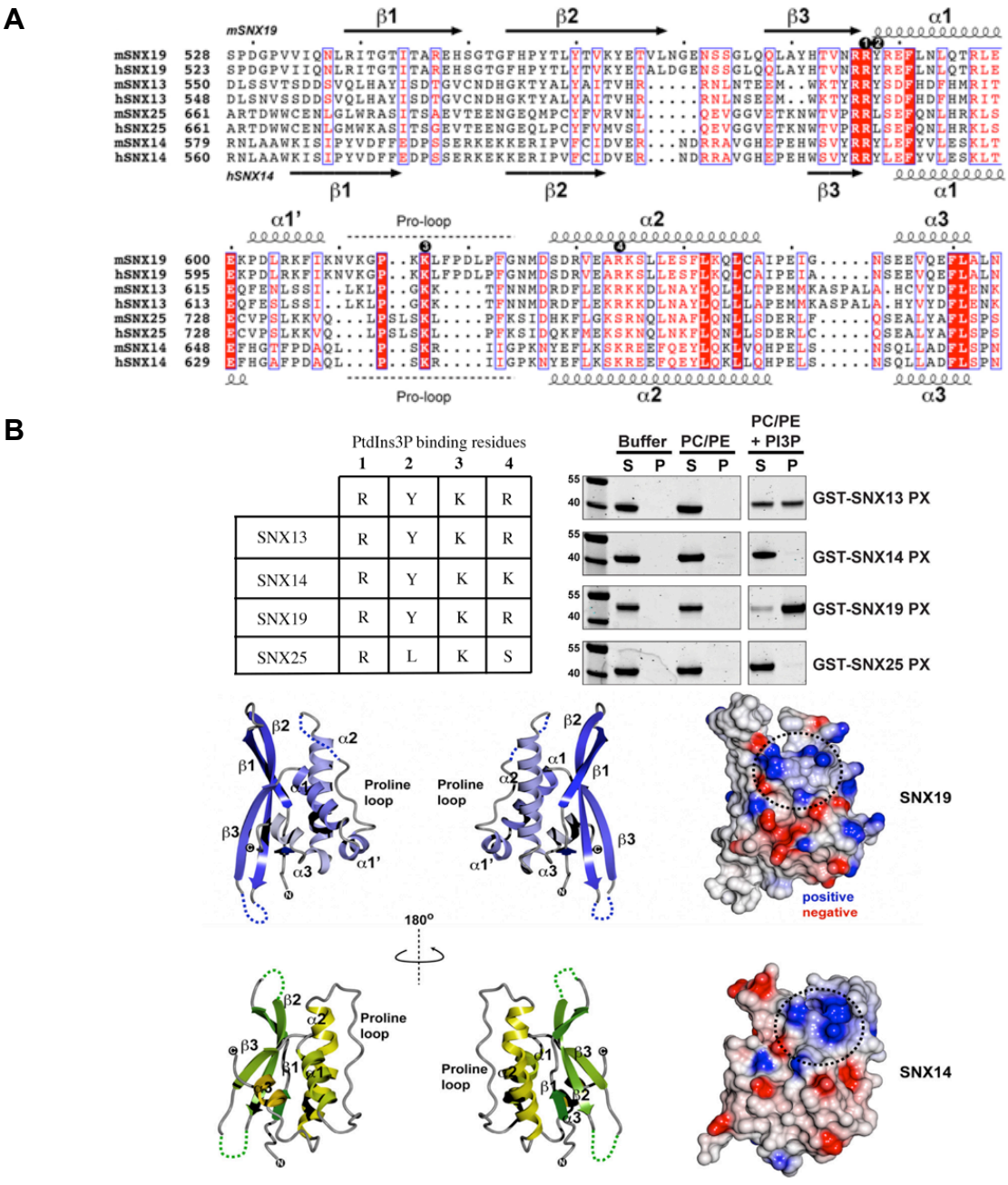


Figure 13. *PX domains of SNX-RGS proteins. (**A**) Amino acid sequence alignment of human and mouse SNX-RGS PX domains showing the PtdIns3P binding residues marked as 1, 2, 3 and 4. (**B**) Supernatant (S) and pellet (P) fractions from the liposome pelleting assay of human SNX-PX PX domain with PtdIns3P, and the X-ray crystal structures of SNX19 (4P2I) and SNX14 (4PQO) PX domains (taken from (38)).*

Previous crystal structures of PX domains of SNX9 (23) and p40phox (24) in complex with PtdIns3*P* reveal the presence of four side chains that are required for PtdIns3*P* binding. They are Arg, Tyr, Lys and Arg residues marked in **Fig. 13A** as 1, 2, 3 and 4 respectively. These residues bind to the 3-phosphate, inositol ring, 1-phosphate and 4-,5-hydroxyls of PtdIns3*P* respectively. SNX13 and SNX19 contain these critical residues for PtdIns3*P* binding, whereas they are not all conserved in SNX14 and SNX19. As seen in **Fig. 13B** SNX14 contains a Lys instead of Arg at site 4, whereas SNX25 possesses a Leu at site 2 instead of Tyr and Ser instead of Arg at site 4. These studies suggested a significantly altered binding pocket in SNX14 and SNX25, which will make these proteins unable to coordinate the 3' phosphorylated phosphoinositide in their binding pocket in the same way as SNX13 and SNX19 (38).

Liposome pelleting assays and NMR studies of phosphoinositide binding confirmed the predictions from SNX-RGS PX domain sequences and X-ray crystallographic structures. SNX13 and SNX19 were found to bind to membranes supplemented with PtdIns3*P*, whereas SNX14 and SNX25 did not show significant binding to this endosomal phosphoinositide (38). Unpublished studies from the Collins lab indicate that while SNX14 does not bind to any phosphoinositides, SNX25 actually binds to a large variety of phospholipids via a novel lipid-binding site distinct from the canonical lipid binding site (C. Mas and B. Collins unpublished data). The functional significance of these divergent phosphoinositide binding properties remains to be explored.

1.5.3 PX-associated A (PXA) and PX-associated C (PXC) domains

The PXA domain sits N-terminal to the RGS and PX domains, whereas the PXC domain lies downstream to the PX domain in the C-terminal region (**Fig. 9**). The PXA domain is around 200 amino acids in length and the secondary structure predictions indicate this domain is α-helical. In SNX13 and SNX25, a short α-helical region called PXA' follows PXA, which is around 50 amino acids long. The PXC domain is around 150 amino acids in length. Secondary structure predictions indicate that this domain is also α-helical, although PXA and PXC domains are not similar to each other in sequence (38). My PhD project focuses primarily on the PXA domain with the goal of understanding the structure and function of these novel domains. Details on the domains and the specific aims and significance of this project will be described towards the end of this chapter.

1.5.4 What we know so far about the functions of SNX-RGS proteins

The SNX-RGS family to-date remains relatively poorly characterized. The first study of these proteins examined SNX13.

SNX13 is highly expressed in heart, and also in skeletal muscle. SNX13 was first identified as a regulatory protein for $G\alpha_s$, and was shown to form a heteromeric complex with both $G\alpha_s$ and Hrs on endosomes and co-operate in the lysosomal targeting of the EGFR (41,48). Thus SNX13 is suggested to have a bifunctional role both in signalling and trafficking. SNX13 knockout mice show embryonic lethality, and the embryo yolk-sac endoderm cells display disrupted endosome morphology thus suggesting a role in endosomal function that is critical for normal development. This also suggests that other SNX-RGS proteins are unable to compensate for SNX13 loss. Li *et al*. (2014), reported that SNX13 deficiency correlates with severe heart failure and thus might have a potential role in the regulation of cardiac function (50).

SNX14 shows a wider pattern of expression with significantly high expression in the nervous system, whereas complete absence of expression in heart and muscle where SNX13 is known to be abundant (49). Recently, Ha *et al*. (2015) demonstrated SNX14 is involved in accelerated internalization and degradation of 5-hydroxytryptamine (5-HT or serotonin) subtype 6 receptor (5-HT6R). The same report demonstrated that SNX14 translocated to the plasma membrane from endosome-like punctate structures on addition of receptor-activating ligand 5-HT.

SNX19 is expressed in a broad range of tissues including stomach, intestine and stronger levels in testis. This protein plays an important role in the formation of cartilage (chondrogenesis), and prevents cartilage degradation in osteoarthritis (38).

SNX25 is highly expressed in lungs, and has been demonstrated to regulate transforming growth factor (TGF) signalling and lysosome trafficking, and degradation of TGF receptors (51). TGFβ signalling pathways are altered in epilepsy, and SNX25 also expresses in the temporal cortex and has been suggested to be associated with temporal lobe epilepsy (52). The closest homologue of SNX25 is the *D. melanogaster* snazarus, which has been identified to play a role in regulating longevity. Snazarus mutants demonstrate increased lifespans (8,53).

1.5.5 The role of SNX-RGS proteins in disease

While the precise cellular function of the SNX-RGS proteins is still unclear, recent studies have shown an important role for SNX14 in neurological development. This was shown by several labs who found that mutations in SNX14 lead to intellectual disability and cerebellar ataxia, with an associated loss of Purkinje cells in affected individuals (54-56). All disease causing SNX14 mutations result in protein truncations and are likely to significantly affect its expression, stability, and overall function. At the cellular level the

primary finding in both patients and zebrafish models (with SNX14 depletion by morpholino) was an accumulation of late endosomes and autophagic structures containing electron dense undigested cellular material (54,55). This suggests a defect in lysosomal maturation that results in defective breakdown through normal cellular processes. The precise mechanism behind this defect however is unknown.

In addition to the disease associated mutations in SNX14 a recent study has also shown that single-nucleotide polymorphisms in SNX25 are correlated with altered life-expectancy in patients with Alzheimer's disease caused by mutations in the presenilin protein (57). Again the mechanism for this is unclear but it is tempting to speculate that changes in SNX25 expression may affect the normal transport of the transmembrane presenilin-containing γ-secretase protein complex, and therefore affect its ability to cleave the amyloid precursor protein leading to amyloid β-peptide production. Further studies will be needed to confirm this.

1.5.6 Are the RGS-SNX proteins a new family of endosomal membrane tethers?

Until recently the only report on the functions of the poorly characterised PXA and PXC domains was that the SNX25 PXA domain was able to localise to EEA1-positive endosomes independently of the other domains and was important in the role of SNX25 in TGF-β receptor degradation (38,51). In 2014, a new study suggested that SNX13 expression levels might be correlated with heart failure (50). This study found that reduction in SNX13 expression facilitated lysosomal degradative sorting of apoptosis repressor with caspase recruitment domain (ARC). This process leads to apoptotic pathway activation, and results in the loss of cardiac cells and thus a decrease in cardiac function (50). Expression of ARC in the heart has cardioprotective functions against development of heart failure by inhibiting intrinsic and extrinsic apoptotic death pathways through inhibiting caspase-8 activity (58). SNX13 is found to bind directly to ARC and acts as a regulator of the interaction between ARC and caspase-8. ARC sorting may be disrupted under reduced SNX13 expression conditions, thereby causing ARC trafficking defects and premature degradation with subsequent caspase-8 over-activation and apoptotic cardiomyocyte death. Domain deletion approaches in this study suggest that the N-terminal PXA, but not PX, RGS and the PXC domains of SNX13 mediates ARC endosomal sorting (50).

In terms of understanding the precise function of the SNX-RGS proteins perhaps the most exciting finding came recently from studies of the yeast homologue Mdm1p (59). Henne *et al*. (2015), reported that Mdm1p deletion lead to the disruption of the normal

contacts between the yeast endoplasmic reticulum (ER) and the vacuole/endosome, and also to a defect in normal sphingolipid metabolism. Conversely, overexpression of Mdm1p led to increased ER-vacuole tethering. This work suggests that the SNX-RGS proteins may act as novel interorganelle tethering proteins that localise to the ER via their N-terminal transmembrane domains, and form contacts with endosomal membranes through the PX domain. A model for this is presented in **Fig. 14**. This might suggest that this protein family may represent a new class of ER-endosome tethers, mediating normal membrane trafficking and lipid transport and metabolism within the endomembrane system.

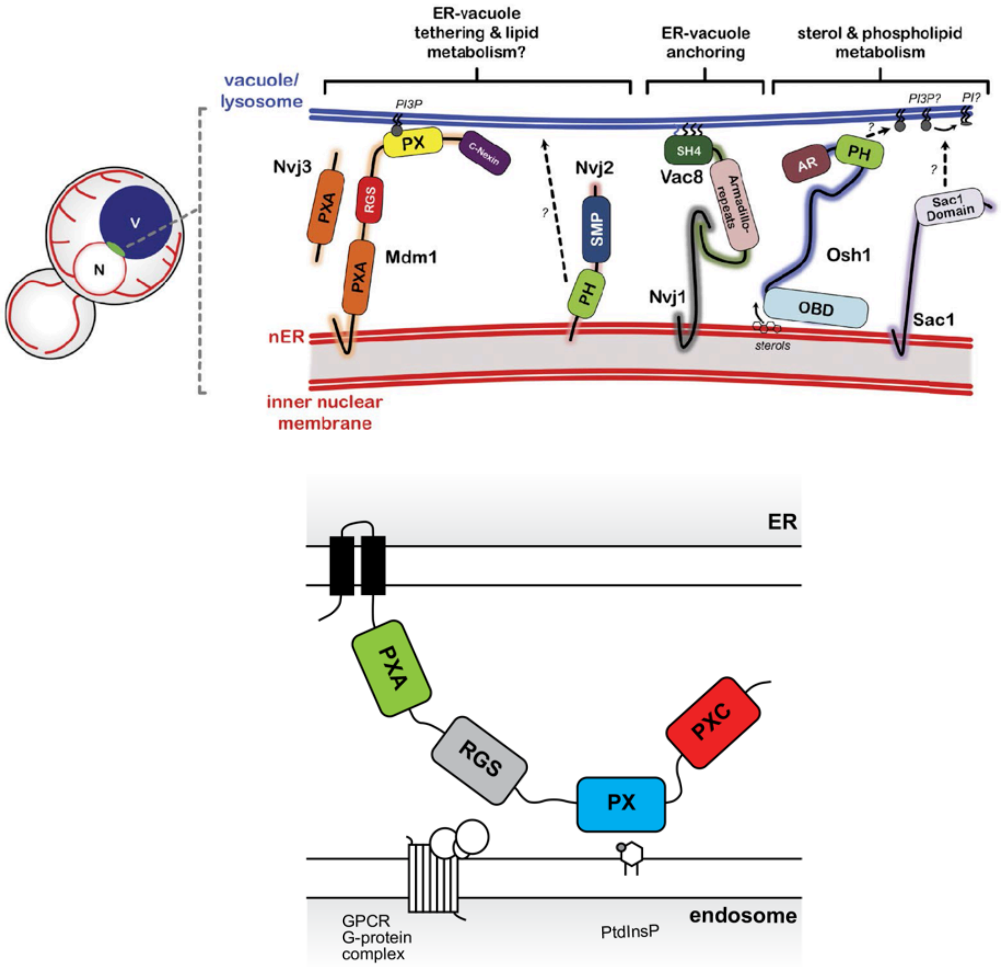


Figure 14. *(**Top**) Model for SNX-RGS proteins showing their potential role in interorganelle tethering and lipid metabolism.* ***(Bottom)*** *SNX-RGS protein model based on the recent findings.*

A more recent study from the same laboratory reports that the yeast ER–vacuole contact sites (NVJs) serve as sites for lipid droplet (LD) biogenesis in response to nutritional stress. Their studies points to the fact that the ER-vacuole tether protein Mdm1p

is sufficient to generate a LD associated ER–vacuole contact site. Mdm1p overexpression expands the NVJ and promotes LD production. It is suggested that the formation of LDs near the vacuole is a metabolic strategy to facilitate their efficient turnover by lipophagy during nutritional starvation. The loss of *MDM1p* may affect the organization or activity of the fatty acid-activating enzymes with which it is found to interact, or even the turnover rate of LDs by lipophagy at the vacuole (60). This recent finding regarding the role of Mdm1p at the ER-vacuole site in yeast may suggest that the mutations in SNX14 leading to neuronal defects in autophagy (54,55) are caused by defects in the ER-tethering process or disruption to the lipid droplet function. Neurological diseases are reported to cause accumulation of sphingolipids and vacuoles in intracellular granular-like deposits (59,61). Thus it is speculated that the disease-associated SNX14 alleles on top of perturbing ER-endolysosomal communication, may perturb lipid droplet biogenesis, contributing to the neurological disease.

1.6 Aims and scope of this thesis

Although the role of the PX domain in association with PtdIns3*P* has been studied in detail, the potential of PX domains to mediate protein-protein interactions is very poorly characterised. In addition, there is also a significant lack of knowledge regarding the structure and function of domains other than PX domains found within proteins of the SNX and PX family (8). A better understanding of the mechanisms of action of these additional domains is essential for understanding their roles in endosomal trafficking and signalling. This will also help us to determine whether SNXs are suitable as therapeutic targets, and if so how such interventions might be achieved. The SNX proteins, as described above, even though they fall under the same larger PX-protein family, play diverse roles in endosomal transport, remodelling and cell signalling. Thus the overall aim of my PhD project is to study the structure and function of SNX proteins from two distinct subfamilies.

(i) SNX-BAR proteins: *SNX5, SNX6 and SNX32, which contain an additional Bin/amphiphysin/Rvs (BAR) domain, display a novel ability to associate with secreted effector proteins of bacterial pathogens.*

By understanding the structural basis for IncE binding of SNX5, SNX6 and SNX32 will enable us to understand how intracellular pathogens like Chlamydia and the proteins involved manipulate the host cell and mammalian retrograde pathway to cause infection. This work also provides the very first insights into how a PX domain binds to another protein rather than lipids. This also provides an idea that the PX-domains of SNX5, SNX6

and SNX32 can potentially regulate the recycling of certain transmembrane cargos.

(ii) SNX-RGS proteins: *SNX13, SNX14, SNX19 and SNX25, which may play an important role in mediating interactions between the membranes of the ER and endosomal compartments, and also potentially act as GAP proteins for $G\alpha_s$ subunits.*

When I began my PhD project, little was known about the precise cellular functions of the SNX-RGS proteins, except that they were highly conserved and are required for normal development. Thus the goal was to use structural biology in order to gain insights into the functions of these proteins, with a particular focus on the structurally uncharacterised PXA and PXC domains. Our initial hypothesis was that they could represent novel protein-protein interaction sites. New findings published in the last 24 months that SNX-RGS proteins play a critical role in neurological disease, and that their function may be to regulate communications between ER, LDs and endosomal compartments are changing the way we view this protein family. In particular, it opens possibilities that the SNX-RGS proteins may not only participate in phosphoinositide and $G\alpha$ interactions through PX and RGS domains but also can control membrane tethering and regulate lipid transfer. Thus my structural studies of the PXA domain will in the future be combined with analysing whether they participate in novel protein-protein interactions *and/or* protein-lipid binding. This project also involves understanding the GAP activity of the RGS domains for $G\alpha_s$ subunits.

In **chapter 2**, (presented as a first author paper published in *eLife*), I provide the first insight into the structural mechanism behind the binding of IncE with the SNX5 PX domain. These findings are confirmed using biophysical techniques and also using cell studies performed in the lab of Assoc. Prof. Rohan Teasdale (UQ). Besides these findings we also shed light into the functional role of the highly conserved binding site on SNX5 PX domain.

In **chapter 3**, (part of an article in the process of submission), we elucidate the high-resolution structure of the SNX32 PX domain in complex with the Chlamydial protein IncE. The work validates the biophysical experiments presented in chapter 2 and also confirms the mode of binding here is the same as SNX5 PX-IncE binding.

In **chapter 4**, (part of an article to be submitted), I build on the novel findings from my studies of the SNX5-IncE interaction. These studies identified a conserved surface groove in SNX5, SNX6 and SNX32 that I speculated might be involved in binding to endogenous proteins. This idea was subsequently supported by work from the lab of our collaborator Prof. Peter Cullen (University of Bristol). In chapter 4, I examine the direct binding of the SNX5/6/32 PX domains to putative cargo molecules identified by proteomics

in the lab of Prof. Peter Cullen. Biophysical experiments reveal peptide sequences conserved in several cargoes bind directly to the PX domains of these SNX-BAR proteins. I also provide crystallographic data presenting the first molecular explanation of how critical endosomal cargos bind to these selected SNX-BAR proteins.

In **chapter 5**, (chapter to be submitted as an article), I further analyse the structural data from chapter 2 and 3, where the structure of IncE bound to SNX5 revealed a tight β-hairpin peptide interaction. I speculated that cyclisation of this IncE sequence might enhance and stabilise the interaction with SNX5PX. ITC experiments of many different variants of IncE-derived cyclic peptides confirm this hypothesis. I discuss how such cyclic peptides could now be used as tools for studying the functions of the SNX-BAR protein family in cargo trafficking, and potentially as leads for therapeutically modulating their activity.

In **chapter 6**, (chapter to be submitted as an article), I systematically investigate the ability of human SNX13, SNX14 and SNX25 RGS domains to bind to the $G\alpha_s$ subunits using *in vitro* GST pull downs and ITC experiments, and their GAP activity against $G\alpha_s$ *in vitro*. I also present the X-ray crystallographic structure of the apo SNX25 RGS domain, and my initial screens of the RGS-$G\alpha_s$ complexes.

In **chapter 7**, I describe my attempts to express and purify the novel PXA domain of SNX-RGS proteins. I was able to design constructs and define purification protocols that eventually yielded soluble and homogeneous protein samples, and identified hits in initial crystallisation screens. Future work will aim to optimize these constructs and crystals for high-resolution structure determination.

Finally, in **chapter 8** I provide an overall summary of the work as well as the future directions.

Chapter 2

Chapter 2

Structural basis for the hijacking of endosomal sorting nexin proteins by *Chlamydia trachomatis*

Blessy Paul[1], Hyun Sung Kim[1], Markus C. Kerr[1], Wilhelmina M. Huston[2], Rohan D. Teasdale[1]* and Brett M. Collins[1]*

1. Institute for Molecular Bioscience, The University of Queensland, St. Lucia, Queensland, 4072, Australia.
2. School of Life Sciences, University of Technology Sydney, Broadway, New South Wales 2007, Australia.

*Correspondence should be addressed R.D.T. (email: r.teasdale@imb.uq.edu.au) and B.M.C (e-mail: b.collins@imb.uq.edu.au).

Running Title: Structure of IncE bound to SNX5

Keywords: endosome, retromer, SNX-BAR, chlamydial inclusion, IncE

This chapter is presented as a published article in *eLife*.

2.1 Abstract

During infection chlamydial pathogens form an intracellular membrane-bound replicative niche termed the inclusion, which is enriched with bacterial transmembrane proteins called Incs. Incs bind and manipulate host cell proteins to promote inclusion expansion and provide camouflage against innate immune responses. Sorting nexin (SNX) proteins that normally function in endosomal membrane trafficking are a major class of inclusion-associated host proteins, and are recruited by IncE/CT116. Crystal structures of the SNX5 phox-homology (PX) domain in complex with IncE define the precise molecular basis for these interactions. The binding site is unique to SNX5 and related family members SNX6 and SNX32. Intriguingly the site is also conserved in SNX5 homologues throughout evolution, suggesting that IncE captures SNX5-related proteins by mimicking a native host protein interaction. These findings thus provide the first mechanistic insights both into how chlamydial Incs hijack host proteins, and how SNX5-related PX domains function as scaffolds in protein complex assembly.

2.2 Introduction

To counter host defence mechanisms intracellular bacterial pathogens have evolved numerous strategies to evade immune detection, replicate and cause infection. Many pathogens manipulate endocytic pathways to gain entry into host cells and generate a membrane-enclosed replicative niche. This frequently involves hijacking or inhibiting the host cell trafficking machinery, first to generate the pathogen containing vacuole (PCV) and subsequently to prevent fusion with lysosomal degradative compartments. Concomitantly the pathogen often endeavors to decorate the PCV with host proteins and lipids that mimic other host cell organelles in order to circumvent innate immune detection, expand the replicative niche and acquire nutrients to support intracellular replication (62,63). This process is orchestrated through the action of molecular syringe-like secretion systems that deliver bacterial effector proteins directly into the host cell cytoplasm.

Chlamydia trachomatis is arguably one of the most successful human bacterial pathogens by virtue of its capacity to hijack host cell intracellular trafficking and lipid transport pathways to promote infection (31,64-66). *C. trachomatis* causes nearly 100 million sexually transmitted infections annually worldwide, and if left unchecked leads to various human diseases including infection-induced blindness, pelvic inflammatory disease, infertility and ectopic pregnancy (67,68). Even though chlamydial infections can generally be treated with antibiotics, persistent infections remain a challenge (69,70).

All Chlamydiae share a common dimorphic life cycle, where the bacteria alternates between the infectious but non-dividing elementary body (EB) form, and the non-infectious but replicative reticulate body (RB) form. Following internalization of EBs through a poorly defined endocytic process, the bacteria reside in a membrane-bound vacuole termed the inclusion, where EBs convert into RBs and replication occurs over 24-72 hours. RBs eventually redifferentiate back to EBs in an asynchronous manner, and are then released to infect neighboring cells (30,62,71). The encapsulating inclusion membrane provides the primary interface between the bacteria and the host cell's cytoplasm and organelles. From the initial stages of invasion until eventual bacterial egress the chlamydial inclusion is extensively modified by insertion of numerous Type-III secreted bacterial effector proteins called inclusion membrane proteins or "Incs". The Incs modulate host trafficking and signaling pathways to promote bacterial survival at different stages, including cell invasion, inclusion membrane remodeling, avoidance of the host cell innate immune defense system, nutrient acquisition and interactions with other host cell organelles (31,66,72).

Chlamydiae secrete more than fifty different Inc proteins. While Incs possess little sequence similarity, they share a common membrane topology with cytoplasmic N- and C-terminal domains, separated by two closely spaced transmembrane regions with a short intra-vacuolar loop (32,72-75) (**Fig. 15**). The cytoplasmic N- and C-terminal sequences of the Inc proteins act to bind and manipulate host cell proteins. Reported examples include the binding of the small GTPase Rab4A by CT229 (76), Rab11A by Cpn0585 (77), SNARE proteins by IncA (78), centrosomal and cytoskeletal proteins by Inc850 and inclusion protein acting on microtubules (IPAM) (79-81), myosin phosphatase by CT228 (82), 14-3-3 and Arf family proteins by IncG and InaC (83,84), and the lipid transfer protein CERT by IncD (85,86). Despite these reports, there are no known structures of Inc family members either alone or in complex with host effectors.

Two recent studies have greatly expanded the repertoire of host cell proteins known to associate with chlamydial inclusions and Inc proteins (26,27). Both reports confirmed that membrane trafficking proteins are major components of the inclusion proteome; and in particular members of the endosomal sorting nexin (SNX) family are highly enriched. Specifically it was shown that the *C. trachomatis* IncE/CT116 protein could recruit SNX proteins containing bin-amphiphysin-Rvs (BAR) domains SNX1, SNX2, SNX5 and SNX6 (26). SNX1 and SNX2 are highly homologous and form heterodimeric assemblies with either SNX5 or SNX6 to promote endosomal membrane tubulation and trafficking (87). A fifth protein SNX32 is highly similar to SNX5 and SNX6 but is almost exclusively expressed in the brain and has not yet been characterized. SNX recruitment to the inclusion occurs via the C-terminal region of IncE interacting with the phox-homology (PX) domains of SNX5 or SNX6 (26) (**Fig. 15**). Interestingly, RNAi-mediated depletion of SNX5/SNX6 does not slow infection but rather increases the production of infectious *C. trachomatis* progeny suggesting that the SNX recruitment is not done to enable bacterial infection. Instead it was proposed that because SNX proteins regulate endocytic and lysosomal degradation, the manipulation by IncE could be an attempt to circumvent SNX-enhanced bacterial destruction (26,27).

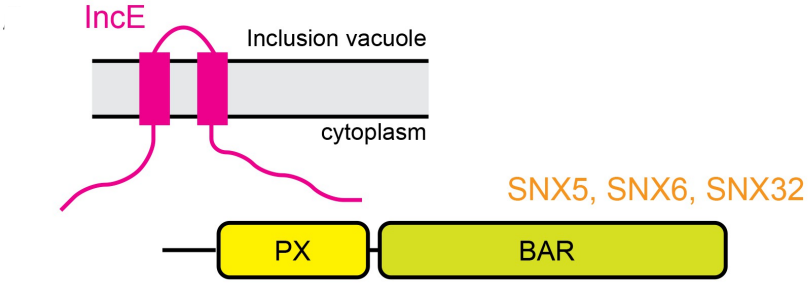


***Figure 15**. IncE from C. trachomatis binds the PX domains of SNX5, SNX6 and SNX32. Schematic depiction of the domain organisation and interactions between IncE and SNX5-related proteins.*

Here we use X-ray crystallographic structure determination to define the molecular mechanism of SNX5-IncE interaction, and confirm this mechanism using mutagenesis both *in vitro* and in cells. When bound to SNX5, IncE adopts an elongated β-hairpin structure, with key hydrophobic residues docked into a complementary binding groove encompassing a helix-turn-helix structural extension that is unique to SNX5, SNX6, and the brain-specific homologue SNX32. A striking degree of evolutionary conservation in the IncE-binding groove suggests that IncE co-opts the SNX5-related molecules by displacing a host protein (as yet unidentified) that normally binds to this site. Our work thus provides both the first mechanistic insights into how protein hijacking is mediated by inclusion membrane proteins, and also sheds light on the functional role of the SNX5-related PX domains as scaffolds for protein complex assembly.

2.3 Materials and methods

2.3.1 Peptides

All synthetic peptides used for isothermal titration were purchased from Genscript (USA). For ITC experiments, peptides were weighed and dissolved in 50 mM Tris (pH 8.0) and 100 mM NaCl (ITC buffer) to make a stock peptide concentration of 2 mM, which was diluted to 0.75 mM before use.

2.3.2 Antibodies and reagents

Polyclonal antibody against *C. trachomatis* HtrA was generated previously (88). Monoclonal antibodies against EEA1 (610457, 1:100) and SNX1 (611483,1:100) were supplied by BD Bioscience. Monoclonal antibodies against the myc epiptope (9B11, 1:2000) were supplied by Abcam. Rabbit polyclonal antibodies against GFP (A-6455, 1:500) were purchased from Molecular Probes (Invitrogen). Secondary antibodies were purchased from Molecular Probes (Life Technologies) and Li-Cor Bioscience. Wortmannin was supplied by Sigma-Aldrich (W1628). VPS34-In1 was from Merck Millipore (532628).

2.3.3 Molecular biology and expression constructs

The IncE sequence used in this study is from the L3 serovar L3/404/LN (NCBI reference WP_015506602) (89). The pGEX-4T-2 bacterial expression plasmid encoding the human SNX5 PX domain (residues 22-170) was generated using a standard PCR-based cloning strategy, and its identity confirmed by sequencing. All other bacterial expression constructs for human SNX proteins were synthesized and cloned into pGEX-4T-2 by Genscript (USA). These included the SNX5 PX domain IncE fusion (SNX5 residues 22-

170 with IncE residues 108-132 fused at the C-terminus, SNX6 (residues 29-170), SNX32 (residues 17-166), and SNX5 PX domain mutants. The pcDNA3.1-N-eGFP mammalian expression constructs encoding full-length human SNX5, SNX5(F136A), IncE(91-132) and IncE(91-132)(F116D) with N-terminal GFP-tags were generated by Genscript (USA). The pCMU-myc-SNX5 was as described previously (90), and the SNX6 and SNX32 genes cloned into the pcDNA3.1-nMyc vector at BamHI and XhoI restriction sites (91). SNX5, SNX32 and SNX8 were also cloned by polymerase chain reaction, restriction digest and ligation into pEGFP-C1 for expression with N-terminal GFP tags as described previously (92).

2.3.4 Recombinant protein expression and purification

All proteins except SNX5 PX domain mutants were expressed in *Escherichia coli* Rosetta cells, whereas mutant constructs were expressed in BL21 Codon Plus supplemented with appropriate antibiotics. Single colonies from cultures grown on LB agar plates were inoculated into 50 mL LB^{2+} with ampicillin (0.1 mg/mL) and chloramphenicol (0.1 mg/mL), and grown at 37°C with shaking overnight. The following day, 30 mL from the overnight culture was used to inoculate 1 L LB media containing ampicillin (0.1 mg/mL) and chloramphenicol (0.1 mg/mL) and incubated at 37°C. Cells were grown to an optical density (OD) of 0.5-0.6 at 600 nm and induced with 0.5 mM isopropyl-β-D-thiogalactopyranoside (IPTG) (except for the SNX5-IncE fusion, where expression was induced at OD_{600} of 0.8 with 1 mM IPTG). Cultures were incubated with shaking overnight at 18°C until the cells reach an O.D of 3.0 (~24 h). Cells were harvested using a Beckman rotor JLA 8.1000 at 4000 RPM for 30 min at 4°C. Pellets were resuspended in 10 mL lysis buffer (50 mM Tris (pH 8.0), 100 mM NaCl, 5% glycerol, 1 mM DTT, 0.1 mg/ml benzamidine, 0.1 mg/ml DNase) per litre of culture. The cells were subjected to cell disruption and centrifugation using JA 25.50 rotor at 18,000 RPM for 30 min at 4°C. The soluble fractions were first purified using affinity chromatography with glutathione-sepharose, and when required the GST tags were cleaved by thrombin while still bound to the column. The proteins were eluted in 50 mM Tris (pH 8.0), 100 mM NaCl, 5% glycerol, and 1 mM DTT, and then further polished using gel filtration chromatography (Superdex 200, GE healthcare) in a buffer containing 50 mM Tris (pH 8.0), 100 mM NaCl. The fractions corresponding to the respective proteins were then pooled and used directly for ITC or were further concentrated for crystallisation.

2.3.5 Isothermal titration calorimetry

ITC experiments were performed on a Microcal iTC200 instrument at 25°C. The proteins were buffer exchanged into ITC buffer (50 mM Tris (pH 8.0) and 100 mM NaCl) by gel filtration prior to ITC experiments. IncE peptides at 750 μM were titrated into 50 μM PX domain samples. The binding data was processed using ORIGIN 7.0 with a single site binding model to determine the stoichiometry (n), the equilibrium association constant K_a ($1/K_d$), and the enthalpy (ΔH). The Gibbs free energy (ΔG) was calculated using the equation $\Delta G = -RTln(K_a)$; binding entropy (ΔS) was calculated by $\Delta G = \Delta H - T\Delta S$. Three experiments were performed for each set of samples to determine the average ± standard error of the mean (SEM) for thermodynamic quantities, except for the peptide truncation experiments where only single experiments were performed. For these truncated peptide experiments, all experiments were performed using the same batch of protein to allow direct comparions to be made.

2.3.6 Crystallisation, data collection and structure determination

The SNX5 PX domain fusion with IncE was concentrated to 15 mg/ml for crystallization. Eight 96-well crystallization hanging-drop screens were set up using a Mosquito Liquid Handling robot (TTP LabTech) at 20°C. Optimized diffraction-quality crystals were obtained using streak seeding in sitting drop vapor diffusion plates. Crystallisation solutions are provided in **Table 3.** Data were collected at the Australian Synchrotron MX1 and MX2 Beamlines, integrated with iMOSFLM (93), and scaled with AIMLESS (94) in the CCP4 suite (95). The structures were initially solved by molecular replacement with PHASER (96) using the apo-SNX5 PX domain crystal structure as the input model (PDB code 3HPB), minus the extended α-helical domain. The resulting model was rebuilt with COOT (97), followed by repeated rounds of refinement with PHENIX (98). All structural figures were generated using PyMOL (DeLano scientific).

2.3.7 Cell culture and transfections

HeLa cells stably expressing mCherry-Rab25 were previously generated within the lab (99) and were maintained in DMEM (Gibco) supplemented with 10% (v/v) Fetal Calf Serum (FCS) (Bovogen) and 2mM L-glutamine (Invitrogen) in a humidified air/atmosphere (5% CO_2) at 37°C. Cells were transfected at 70% confluence with pcDNA3.1-N-eGFP plasmid constructs using Lipofectamine 2000 as per manufacturer's protocol (Invitrogen) and examined 18-24 hrs later. The HeLa cell line used in this study was from America Type Culture Collection (#ATCC CCL2). Parental and stable cells lines were negative for

mycoplasma by DAPI staining, and authenticated by STR profiling (Cell Bank Australia). For inhibitor treatments, cells were treated with either 100 nM wortmannin or 1 μM Vps34-IN1 for 1 h.

2.3.8 Chlamydial infection assays

C. trachomatis serovar L2 (ATCC VR-902B) was used to infect cells at a multiplicity of infection (MOI) of ~0.5. Cells were infected 2 h post-transfection in normal DMEM (Gibco) supplemented with 10% (v/v) FCS (Bovogen) and 2 mM L-glutamine (Invitrogen) in a humidified air/atmosphere (5% CO_2) incubator at 37°C. After 2 h media was replaced with fresh media.

2.3.9 Microscopy

Transfected and infected cells (18-24 h post-infection) were fixed with 4% paraformaldehyde, permeabilised using TritonX-100 (Sigma) and immunolabeled as described previously (99) and counter-stained with DAPI. The coverslips were imaged using a confocal laser-scanning microscope (LSM 710 meta, Zeiss) with 63x oil immersion objective. Time-lapse videomicroscopy was carried out on individual live cells using a Nikon Ti-E inverted deconvolution microscope using a 40x, 0.9 Plan Apo DIC objective, a Hamamatsu Flash 4.0 4Mp sCMOS monochrome camera and 37°C incubated chamber with 5% CO_2. GFP was excited with a 485/20nm LED and captured using a 525/30nm emission filter, and mCherry was excited using a 560/25nm LED and captured using a 607/36nm emission filter. Data was processed using ImageJ (https://imagej.nih.gov/ij/) and compiled using Adobe Illustrator CS6.

2.3.10 Image quantification

The immunofluorescence colocalisation of GFP-SNX5 with chlamydial inclusion membranes imaged on a confocal microscope was measured by Mander's correlation coefficient of red pixel (EEA1 or mCherry-Rab25) over green pixel (GFP-SNX5) signals, which were determined using ImageJ (https://imagej.nih.gov/ij/) with the JACoP plugin (100). A total of 10 cells from two biological replicates were used in the analysis of statistical significance, which was determined using unpaired non-parametric t-Test with Mann-Whitney test using Prism7 software. Punctate structures were automatically counted using ImageJ analyse particle tool across total of 10 cells from two biological replicates. Statistical significance was obtained as stated. To quantify the effect of PI3K inhibitors on SNX recruitment, Z-stacks were captured with a Zeiss 710 confocal laser scanning microscope using a 40x objective. Maximum projections were generated with FIJI

(https://fiji.sc/) and Pearson's correlation coefficients for individual cells determined using the FIJI 'Coloc 2' function with Costes threshold regression and 100 Costes randomisations. Co-localization analyses were conducted on 2 independent experiments from 5 images per condition each containing at least 20 cells (>100 cells analysed per condition).

2.3.11 Co-precipitation of GFP-SNX5 and endogenous SNX1

HeLa cells were transfected with pcDNA3.1-N-eGFP plasmid constructs overnight at 70% confluence and the cells were lysed using lysis buffer (H_2O, 50 mM HEPES, 150 mM NaCl, 1% Triton-X100, 10 mM $Na_4P_2O_7$, 30 mM NaF, 2 mM Na_3VO_4, 10 mM EDTA, 0.5 mM AEBSF and protease inhibitor cocktail). Cell lysates were incubated with GFP nano-trap agarose beads (Protein Expression Facility, UQ) after preclear using protein G-agarose beads (Invitrogen). Protein complexes attached to the beads were detached by boiling for 5 min with 5x denaturing and reducing buffer (0.625 M Tris pH 6.8, 50% glycerol, 10% SDS, 0.25% Bromophenol blue and 500 mM DTT). Denatured and reduced proteins were separated by molecular mass using SDS-PAGE. Proteins were transferred onto PVDF-FL membrane (Immobilon) and were detected by immunoblotting with polyclonal anti-GFP and monoclonal SNX1 antibodies, and near-infrared fluorescent dyes (LI-COR). Immunolabelled proteins were visualised using LI-COR Odyssey imaging system.

2.3.12 Modelling of the SNX5-SNX1 heterodimer

Human SNX5 and SNX1 sequences were submitted to the PHYRE2 server for automated homology-based model building (101). For both proteins the top scoring modelling template was the crystal structure of the SNX9 PX-BAR domains (PDB ID 2RAJ) (23) with Confidence Scores of 100% (and sequence identities of 19% and 16% respectively). The PX domain of the SNX5 model generated using this structural template was missing the extended α-helical insert, so to complete the model the SNX5 PX domain-IncE complex was substituted and a dimer of SNX5 and SNX1 PX-BAR domains generated by overlaying with the SNX9 dimer in the PtdIns3*P*-bound state (PDB ID 2RAK) (23). The resulting model was subjected to simple energy minimisation in PHENIX (98). Conservation of surface residues was computed using the CONSURF server (102).

2.4 Results

2.4.1 IncE specifically binds and recruits SNX5, SNX6 and SNX32 to *C. trachomatis* inclusions

It was previously shown that the sorting nexins SNX1, SNX2, SNX5 and SNX6 are recruited to the inclusion membrane in *C. trachomatis* infected cells (26,27). We first confirmed this for SNX1, SNX2 and SNX5 in HeLa cells infected with *C. trachomatis* serovar L2 (MOI~0.5) for 18 hr. All three proteins were recruited to the inclusion membrane as assessed by co-localisation with the inclusion marker mCherry-Rab25 (**Fig. 16A**) (99), as were GFP-tagged SNX1 and SNX5 but not the more distantly homologous GFP-SNX8 (**Fig. 16B**).

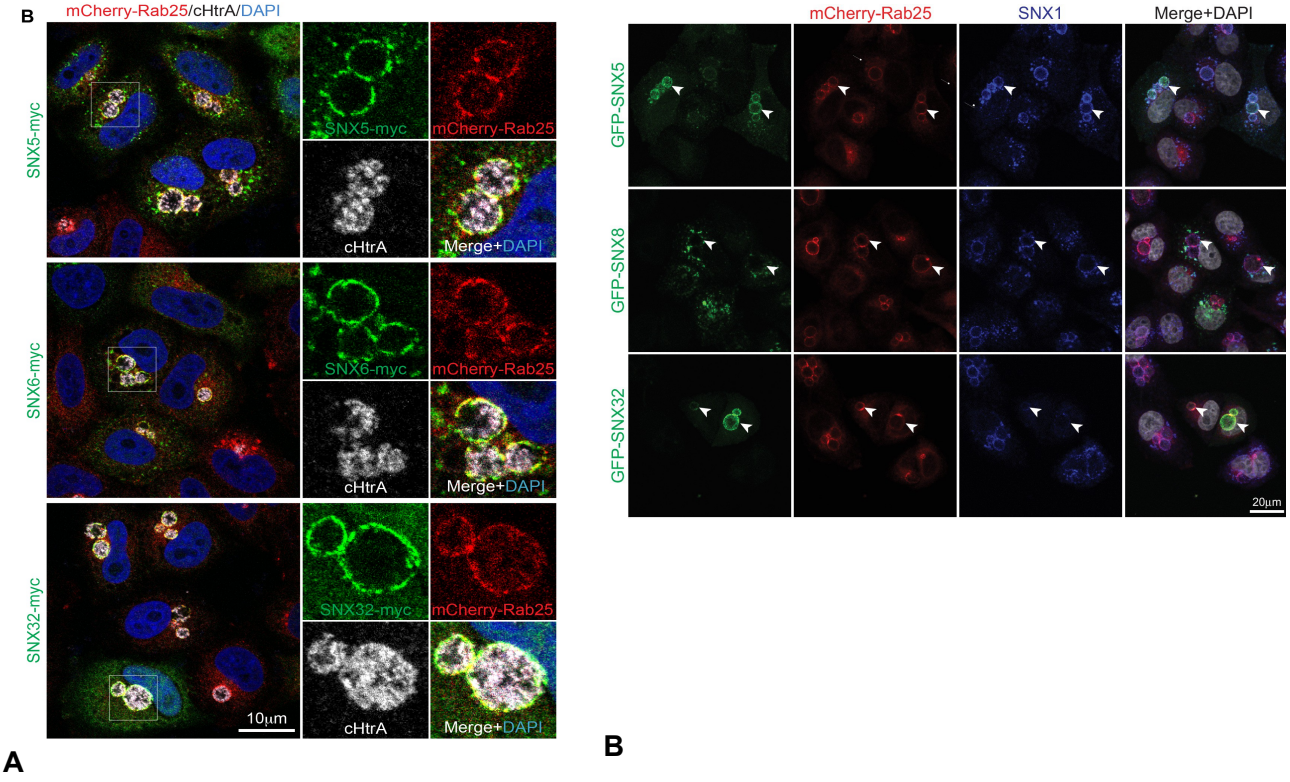


Figure 16. *SNX5, SNX32 and SNX1 are recruited to C. trachomatis inclusions and membrane tubules. (A) HeLa cells stably expressing the mCherry-Rab25 inclusion membrane marker (red) were infected with C. trachomatis serovar L2 (24 hr) and transfected with myc-tagged SNX expression constructs. The samples were fixed and immunolabeled with anti- myc (green) and anti-chlamydial HtrA antibodies (white) and counterstained with DAPI (blue). (B) Hela cells were transiently transfected with GFP-tagged SNX and mCherry-Rab25 proteins as indicated, and infected with C. trachomatis serovar L2. Cells were imaged by confocal fluorescence microscopy for GFP-tagged proteins (green), endogenous SNX1 (blue), mCherry-Rab25 (red) and DAPI-stained nuclear material (white). Both GFP-SNX5 and GFP-SNX32 are recruited to inclusion membranes, but the distantly related SNX-BAR protein SNX8 is not.*

We also observed localization of the SNX proteins to extensive inclusion-associated membrane tubules in a subset of infected cells as described previously (**Fig. 17**) (26,27).

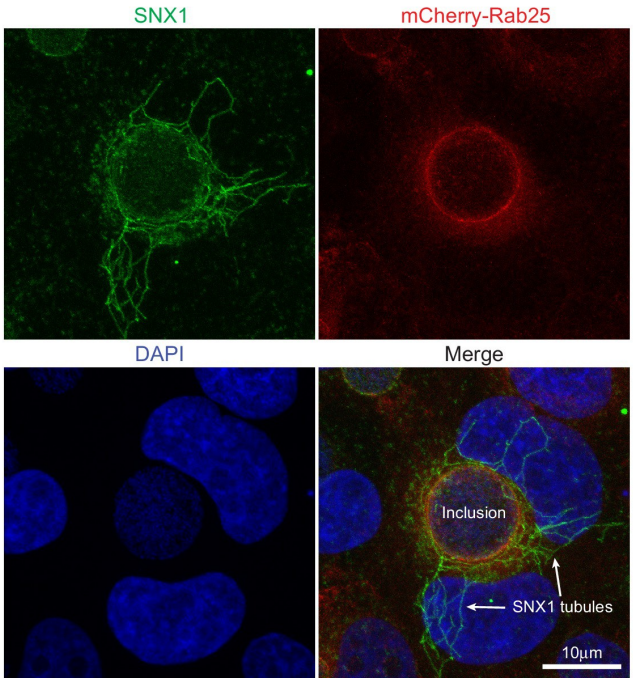


Figure 17*. An example of SNX1-decorated tubules (green) often observed emanating from inclusion membranes (mCherry-Rab25 in red; DAPI staining in blue). The image is a maximum projection.*

Interestingly, when infected cells are treated with wortmannin, a pan-specific inhibitor of phosphoinositide-3-kinase (PI3K) activity, we see a loss of the SNX proteins from punctate endosomes, but not from the inclusion membrane (**Fig. 18; Video 1**). A similar result is seen for specific inhibition of PtdIns3*P* production by Vps34 using Vps34-IN1 (**Fig. 18**). This offers two possibilities; that either SNX recruitment to the inclusion occurs via protein-protein interactions, and does not depend on the presence of 3-phosphoinositide lipids that typically recruit SNX proteins to endosomal membranes, or alternatively that PI3Ks are not present at the inclusion and therefore wortmannin treatment has no effect at this particular compartment. Given our structural and mutagenesis studies below we favor the former explanation.

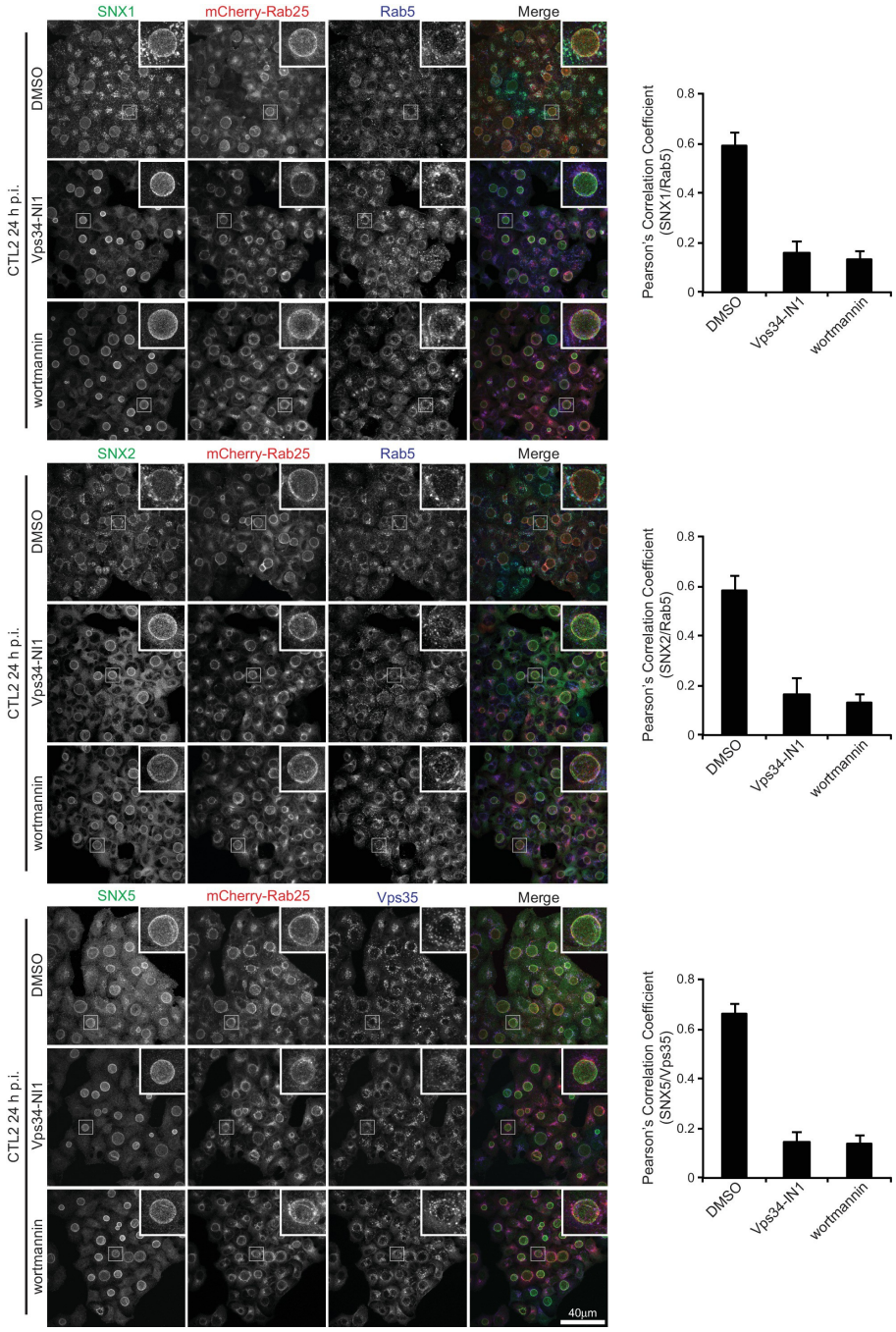


Figure 18. *Recruitment of SNX1, SNX2 and SNX5 to inclusions is not dependent on 3-phosphoinositides. HeLa cells stably expressing mCherry-Rab25 were infected with C. trachomatis serovar L2 (MOI ~0.5) for 24 hr and imaged by immunofluorescence microscopy using antibodies to SNX1, SNX2 and SNX5. mCherry-Rab25 provides marker for the inclusion membrane. The upper panels show control infections and lower panels show cells treated with wortmannin or Vps34-IN1 with concomitant loss of SNX association with endosomal compartments while inclusion localisation is unaffected. The images are maximum projections. Endosomal compartments are labeled with antibodies to endogenous Rab5 or Vps35 and Pearson's correlation coefficients used to quantify loss of endosomal recruitment (100 cells per group; error bars, S.D). A movie showing the effect of wortmannin on GFP-SNX5 is shown in* ***Video 1****.*

Mirrashidi *et al*., (2015), demonstrated an *in vitro* interaction between IncE and the SNX5 and SNX6 PX domains. To confirm their direct association we assessed the binding affinities using isothermal titration calorimetry (ITC) (**Fig. 19A; Table 1**). Initial experiments with the human SNX5 and SNX6 PX domains showed a robust interaction with the IncE C-terminal domain (residues 107-132). The affinities (K_d) for SNX5 and SNX6 were essentially indistinguishable (0.9 and 1.1 μM respectively), but we detected no interaction with the PX domain of SNX1 confirming the binding specificity. The PX domains of SNX5 and SNX6 possess a helix-turn-helix structural insert (25), which is not found in any other SNX family members except for SNX32 (**Fig. 19B**), a homologue expressed primarily in neurons (BioGPS (103)). Confirming a common recruitment motif in the SNX5-related proteins, ITC showed a strong interaction between IncE and the SNX32 PX domain similar to SNX5 and SNX6 (K_d = 1.0 μM) (**Fig. 19A; Table 1**), and SNX32 was robustly recruited to inclusion membranes in infected cells (**Fig. 16A; Fig. 16B**). Overall, our data indicates that a common structure within the SNX5, SNX6 and SNX32 PX domains is required for IncE interaction.

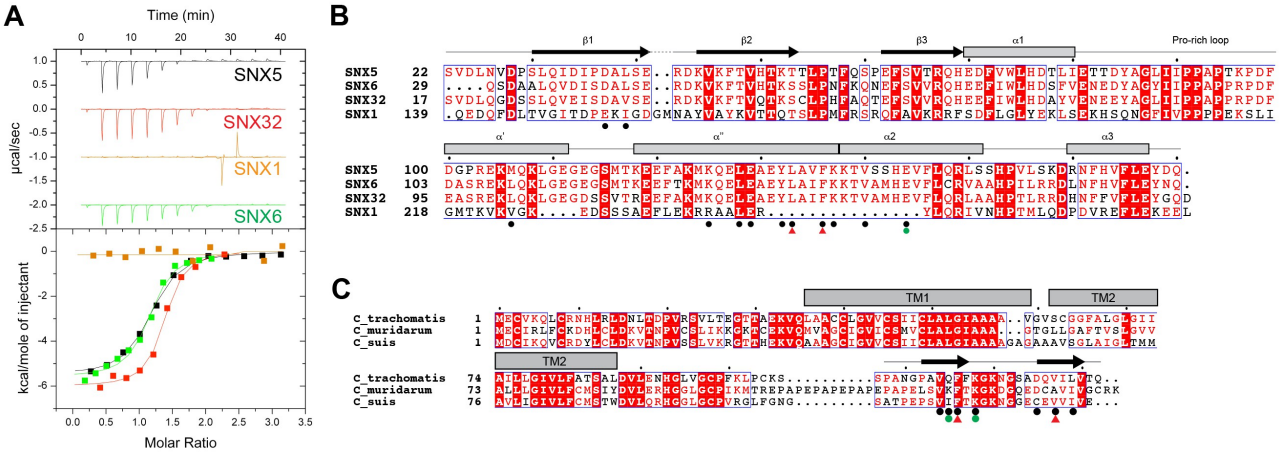


Figure 19. ***(A****) Binding affinity between IncE peptide (residues 107–132) and SNX PX domains by ITC. Top panels show raw data and lower panels show normalised integrated data. (****B****) Sequence alignment of human SNX1, SNX5, SNX6 and SNX32 PX domains. Conserved residues are indicated in red. Side-chains that directly contact IncE in the crystal structure are indicated with black circles. Mutations that block IncE binding are highlighted with red triangles, and mutations that do not affect binding indicated with green circles. Secondary structure elements derived from the SNX5 crystal structure are indicated above. (****C****) Sequence alignment of IncE from C. trachomatis and putative homologues from C. muridarum and C. suis. IncE side-chains that directly contact SNX5 in the crystal structure are indicated with black circles. Mutations that block SNX5 binding are highlighted with red triangles, and mutations that do not affect binding indicated with green circles.*

Finally we tested a series of IncE truncation mutants for their binding to the SNX5 PX domain (**Fig. 20A, 20B and 20C; Table 2**).

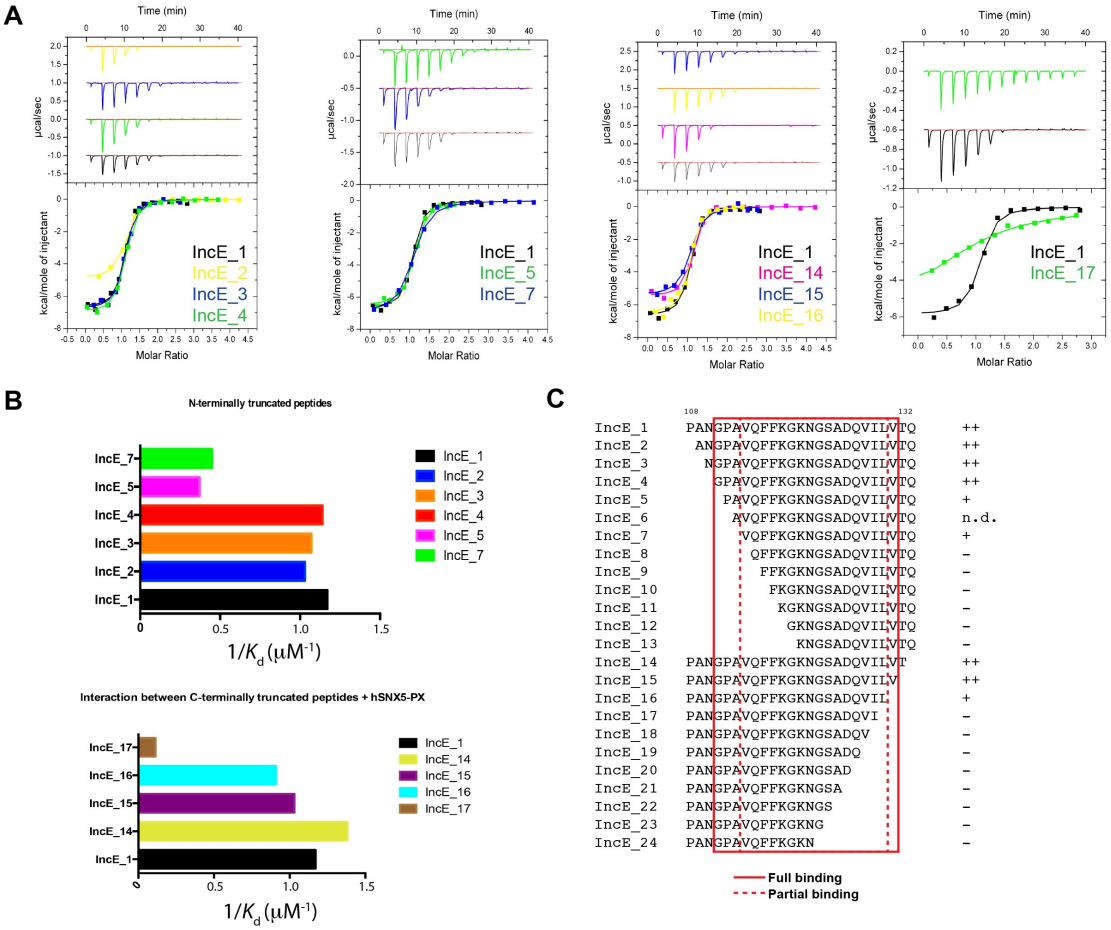


Figure 20. *(**A**) Representative ITC experiments for truncated IncE peptides. These experiments were conducted using a single batch of SNX5 PX domain to minimize batch-to-batch protein variation. (**B**) Plots of the affinity constants for selected peptides to highlight the progressive loss of binding with N and C-terminal truncations. (**C**) Sequences of the truncated IncE peptides are given, with a qualitative indication of binding strength relative to the IncE_1 peptide containing residues 107–132. Full binding is indicated by '++' reduced binding by '+' and lack of binding by '−'.All sequence information and their K_d values are given in **Table 2**. When compared to the reference ITC experiment the binding affinity of peptides was unaffected when the first three N-terminal residues were removed (IncE_2, IncE_3 and IncE_4) and gradually became weaker until IncE_7, after which binding was abolished. Results from IncE_6 are inconclusive due to the difficulty in successfully dissolving the peptides in buffer (n.d.). C-terminal truncations showed that IncE_14 and IncE_15 had similar high binding affinities to the reference, while the binding of IncE_16 and IncE_17 became progressively weaker and peptides shorter than IncE_17 showed no binding. This data indicates that the minimal IncE binding sequence retaining full SNX5 binding is GPAVQFFKGKNGSADQVILVT, and a shorter fragment VQFFKGKNGSADQVIL can bind to SNX5, albeit with a slightly reduced affinity.*

Synthetic peptides were used with single amino acids removed sequentially from the N and C-terminus to determine the minimal sequence required for binding. These experiments showed that the shortest region of IncE able to support full binding to SNX5 encompasses residues 110-131 (GPAVQFFKGKNGSADQVILVT), while a shorter fragment containing residues 113-130 (VQFFKGKNGSADQVILV) can bind to SNX5 with a slightly reduced affinity. While variations are observed across the different *C. trachomatis* serovars (89) the SNX5-binding sequence appears to be preserved in all detected variants (**Fig. 21**). A comparison with other chlamydial species suggests that IncE is not very widely conserved in this Genus, being clearly identifiable only in the closely related mouse-specific *C. muridarum* and swine-specific *C. suis* (**Fig. 19C**). Residues required for binding to SNX5 are preserved in these IncE homologues, but whether SNX proteins are also recruited during infection by these other chlamydial species remains to be determined.

Figure 21. *Sequence alignment of IncE from different C. trachomatis serovars.*

2.4.2 The crystal structure of IncE in complex with the SNX5 PX domain

The canonical PX domain structure is composed of a three-stranded β-sheet (β1, β2 and β3) followed by three close-packed α-helices. The first and second α-helices are connected by an extended proline-rich sequence. Typically PX domains have been found to bind to the endosome-enriched lipid phosphatidylinositol-3-phosphate (PtdIns3*P*) via a basic pocket formed at the junction between the β3 strand, α1 helix and Pro-rich loop. In contrast SNX5, SNX6 and SNX32 possess major alterations in the PtdIns3*P*-binding pocket that preclude canonical lipid head-group docking. In addition they possess a unique extended helix-turn-helix insert between the Pro-rich loop and α2 helix of unknown function (**Fig. 19B**) (25).

To determine the structure of the SNX5-IncE complex we generated a fusion protein encoding the human SNX5 PX domain (residues 22-170) and *C. trachomatis* IncE

C-terminal sequence (residues 108-132) attached at the PX domain C-terminus (**Fig. 22A**). This construct readily crystallised in several crystal forms, and we were able to determine the structure of the complex in three different spacegroups (**Table 3; Fig. 22B**). Confirming that the fusion does not alter complex formation, the short linker region is disordered, and the mode of IncE-binding to SNX5 is identical in all three structures (**Fig. 22C and 22D**).

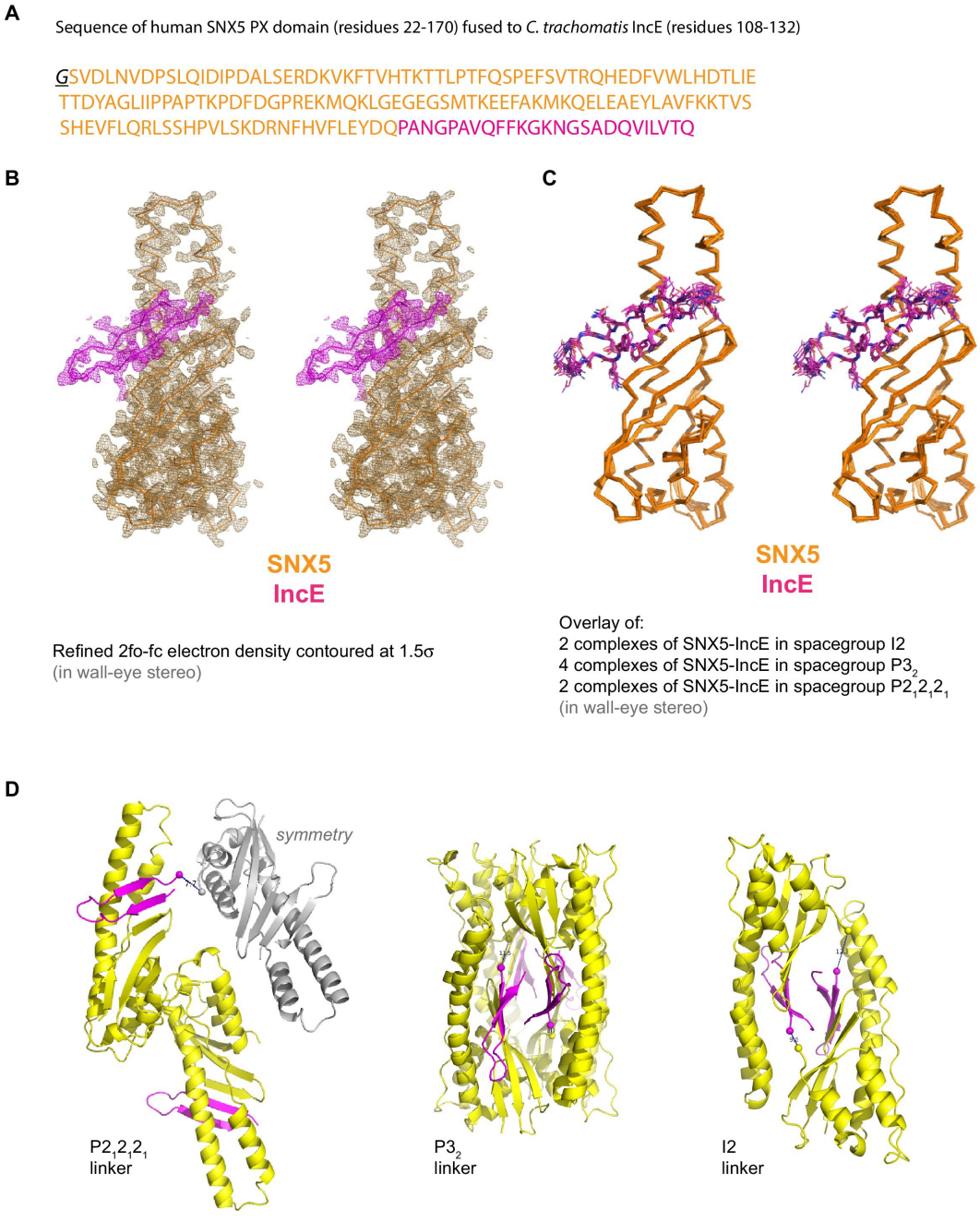


Figure 22. (A) *Sequence of the SNX5 PX domain fusion protein with the IncE C-terminal peptide.* ***(B)*** *Refined 2fo-fc electron density contoured at 1.5σ for the SNX5-IncE structure in spacegroup P212121.* ***(C)*** *Overlay of each independent SNX5-IncE complex observed in the crystal forms.* ***(D)*** *Ribbon structures indicating the locations of the linker regions in each crystal form. The C-terminal SNX5 residues and the N-terminal IncE residues are shown by sphere with distances indicated.*

Because of the higher resolution, we limit our discussions to the structure of the SNX5 PX-IncE complex observed in the $P2_12_12_1$ crystal form. The first three IncE N-terminal residues (Pro107, Ala108, Asn109) and the last three IncE C-terminal residues (Val130, Thr131, Gln132) were not modeled due to lack of electron density, suggesting disorder and matching precisely with our mapping experiments showing these residues are not necessary for SNX5 association.

The IncE sequence forms a long β-hairpin structure that binds within a complementary groove at the base of the extended α-helical insertion of the SNX5 PX domain and adjacent to the β-sheet sub-domain (**Fig. 23A; Video 2**). The β-hairpin structure of IncE (N-terminal βA and C-terminal βB strands) is directly incorporated as a β-sheet augmentation of the β1, β2 and β3 strands of SNX5 (**Fig. 23B**).

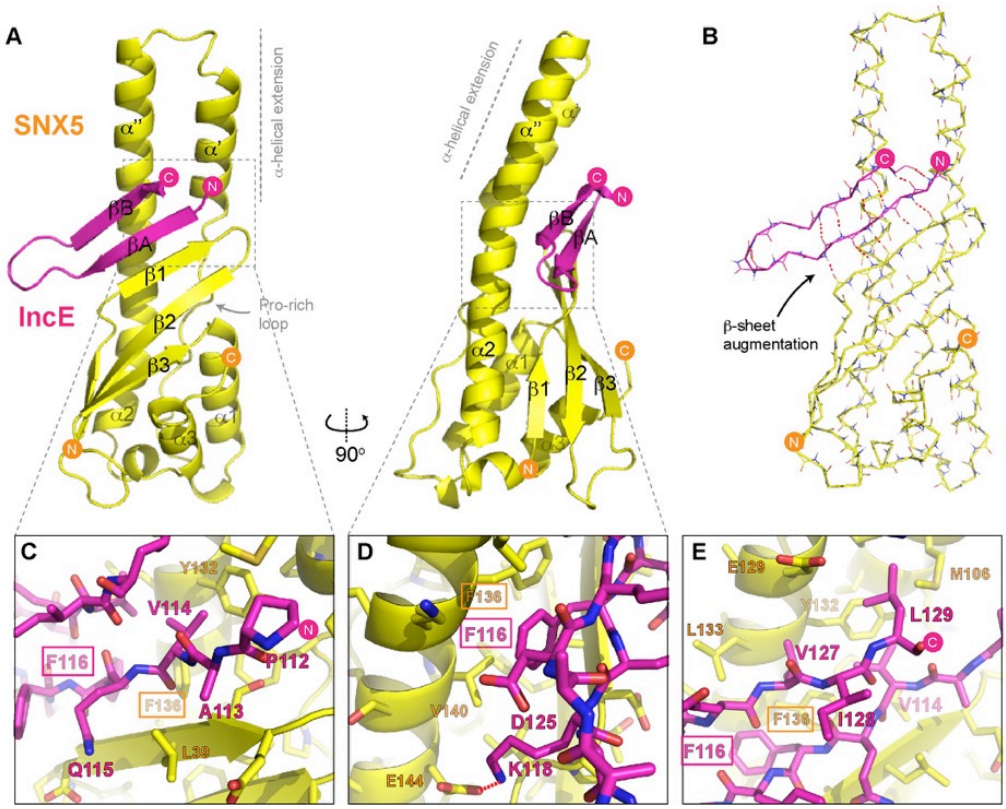


Figure 23. *(**A**) Crystal structure of the SNX5 PX domain (yellow) in complex with IncE residues 107-132 (magenta) shown in cartoon representation. (**B**) Backbone atoms of the SNX5 and IncE proteins are shown to highlight the prominent β-sheet augmentation mediating the association between the two molecules. (**C**) Close up view of the SNX5-IncE interface highlighting specific contact areas at the N-terminus of the IncE peptide. (**D**) Close up of the SNX5-IncE interface highlighting specific contact areas at the hairpin loop of the IncE peptide shown at 90° to **Fig. 23C**. (**E**). Close up of the SNX5-IncE interface highlighting contact areas at the C-terminus of the IncE peptide in approximately the same orientation as **Fig. 23C**. Residues in SNX5 (Phe136) and IncE (Phe116) that are critical for binding based on mutagenesis are boxed.*

The N-terminal βA strand of the IncE sequence (Gly111-Lys118) forms the primary interface with SNX5, making main-chain hydrogen bonds with the β1 strand of the SNX5 PX domain for the stable positioning of the IncE structure. The two anti-parallel β-strands of IncE are connected by a short loop (Gly119-Ala124) that makes no direct contact with the SNX5 protein, and the C-terminal IncE βB strand (Asp125-Val130) forms an interface with the extended α-helical region of the SNX5 PX domain.

Detailed views of the SNX5-IncE interface are shown in **Figs. 23C, 23D and 23E**. Aside from main-chain hydrogen bonding to form the extended β-sheet, IncE engages in several critical side-chain interactions with the relatively hydrophobic SNX5 binding groove. At the N-terminus of the βA strand Val114 of IncE inserts into a pocket formed primarily by Tyr132 and Phe136 on the SNX5 α'' helix (**Fig. 23C**). A major contribution comes from IncE Phe116, with π-stacking occurring with the Phe136 side-chain and hydrophobic docking with Val140 of SNX5 (**Fig. 23D**). Adjacent to IncE Phe116 at the end of the βA strand Lys118 makes an electrostatic contact with SNX5 Glu144. Finally, at the C-terminal end of the IncE βB strand Val127 and Leu129 contact an extended SNX5 surface composed of Leu133, Tyr132 and Met106 (**Fig. 23E**).

2.4.3 Mutations in the SNX5-IncE interface disrupt complex formation *in vitro* and in cells

To verify the crystal structure we mutated residues from both SNX5 and IncE and measured their affinities using ITC (**Fig. 24 top; Table 2**). At the interface between SNX5 and IncE several side chains make key contributions to peptide recognition. Because Leu133 and Phe136 residues in SNX5 are located at the core of the IncE-binding interface, and also due to the structural rearrangements these residues make on binding (see below), we reasoned that L133D and F136A mutations would inhibit the interaction. Indeed these mutants abolished association with the IncE peptide (**Fig. 24A**). The reciprocal mutations in IncE residues F116A and V127D also abolished binding to the SNX5 PX domain (**Fig. 24B**), and the SNX6 and SNX32 PX domains (**Fig. 24C**), demonstrating the importance of these hydrophobic and π-stacking interactions for stable complex formation. In contrast mutations predicted to disrupt an observed electrostatic contact (IncE K118A or SNX5 E144A) had little effect on binding. Thus the core hydrophobic interactions are critical for IncE binding but the peripheral electrostatic contact is not essential.

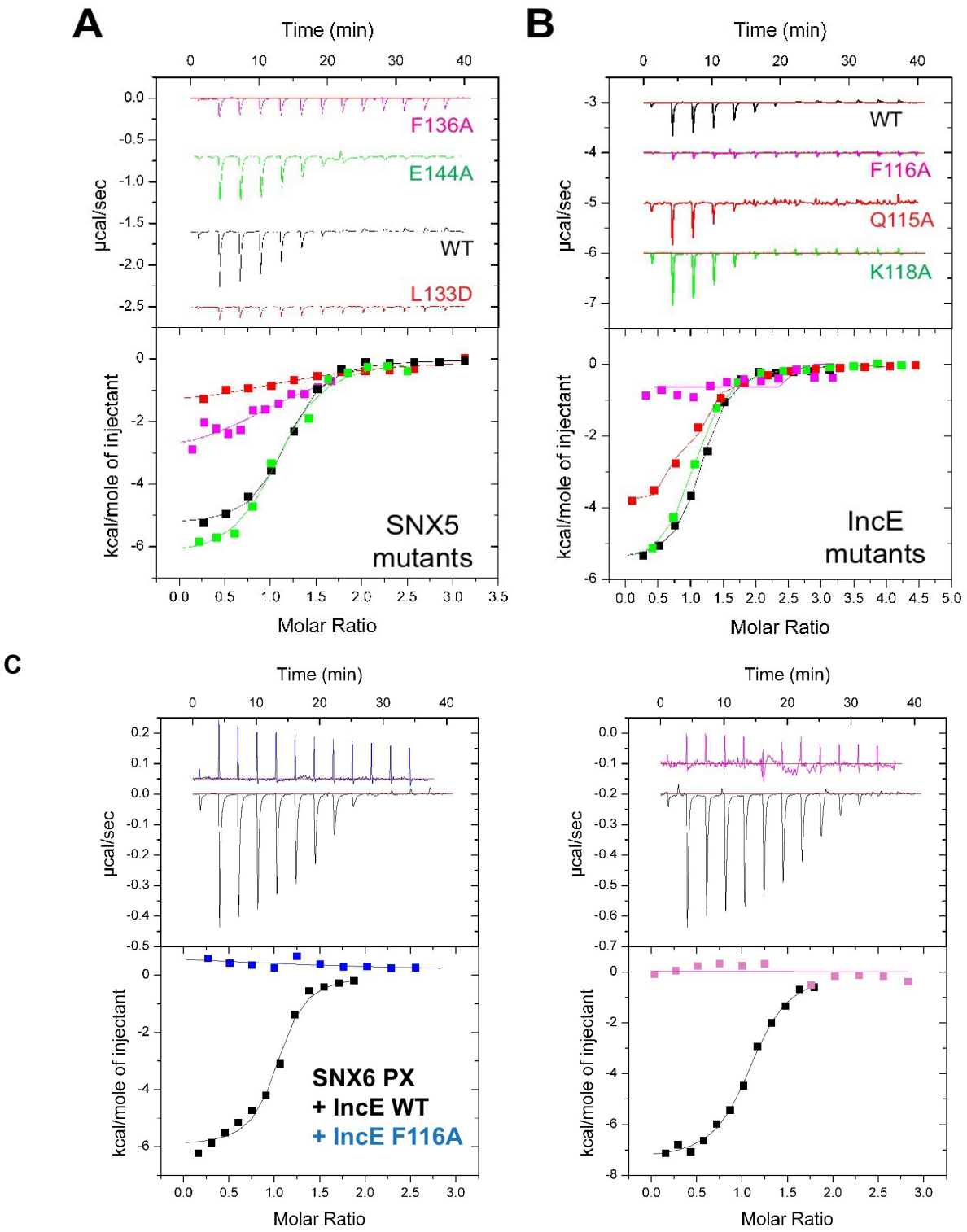


Figure 24. *(Top) Mutations in the SNX5 and IncE proteins prevent complex formation in vitro and in cells (**A**) ITC experiments testing the effect of SNX5 mutations on IncE binding. Both L133D and F136A mutations prevented IncE binding, but the E144A mutation had little effect. (**B**) ITC experiments testing the effect of IncE mutations on SNX5 binding. F116A blocked SNX5 interaction, while Q115A had a partial effect and K118A had no effect on association. **(C)** SNX6 and SNX32 PX domains bind IncE at the same site as SNX5. ITC experiments testing the effect of IncE peptide mutations on binding to SNX6 and SNX32 PX domains. The IncE F116A mutation blocks interaction with both PX domains similarly to SNX5.*

To confirm the role of IncE in direct SNX5 protein recruitment to the chlamydial inclusion we examined the localisation of GFP-tagged SNX5 in HeLa cells infected with *Chlamydia trachomatis* L2 (CTL2) for 24 h (MOI~0.5). Cells expressing the GFP-SNX5 protein showed clear and uniform recruitment to the limiting membrane of the inclusion as defined by mCherry-Rab25 (**Fig. 25A**), which is consistent with the localisation observed by others (26,27). In contrast, the GFP-SNX5 (F136A) mutant protein showed no recruitment to the chlamydial inclusion. The change in relative distribution of these GFP-SNX5 proteins on the inclusion was quantified for wildtype SNX5 (Mander's coefficient 0.67 ± 0.14) and GFP-SNX5 (F136A) (0.041 ± 0.051) (**Fig. 26A**).

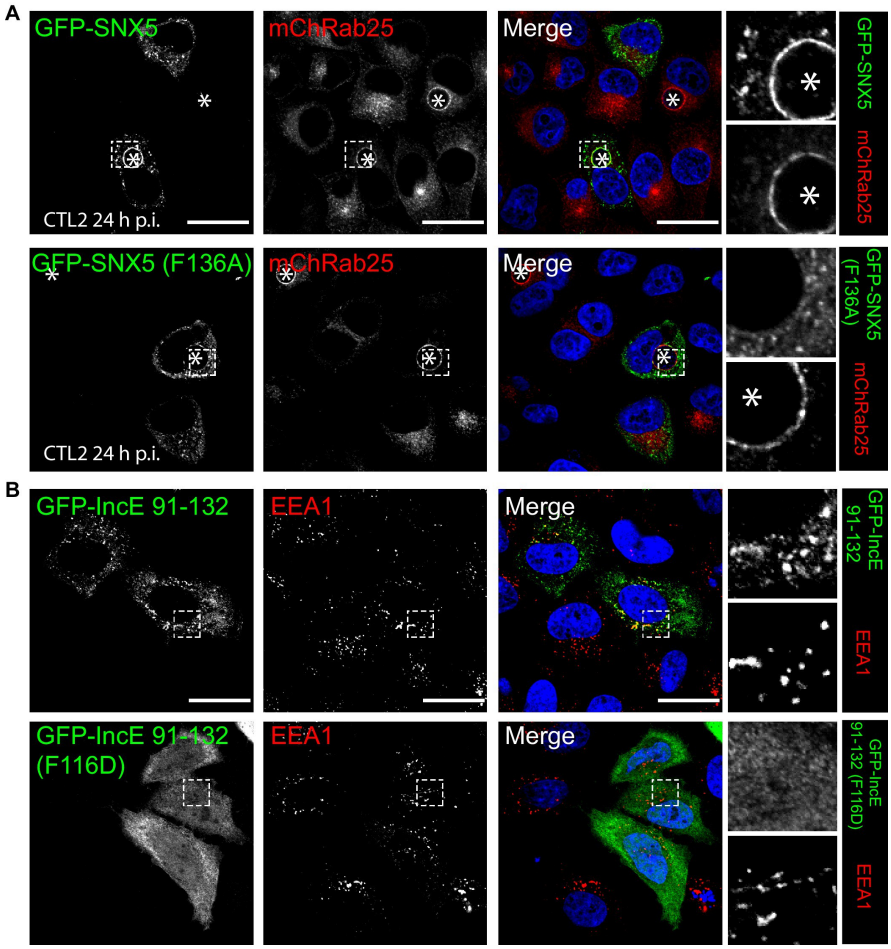


Figure 25. *Mutations in the SNX5 and IncE proteins prevent complex formation in cells (**A**) Single amino-acid mutation in the PX domain of the SNX5 (F136A) abolishes recruitment to the chlamydial inclusion. HeLa cells stably expressing mCherry-Rab25 (red) were transfected transiently with GFP-SNX5 or GFP-SNX5 (F136A) (green) and infected with Chlamydia trachomatis L2 for 18-24 h. The cells were fixed and the nucleic materials were counter-stained with DAPI (blue). (**B**) HeLa cells were transfected transiently with GFP-IncE(91-132) or GFP-IncE(91-132)(F116D) (green) and co-labelled for the early endosomal marker EEA1 (red). Mutation in the SNX5 binding IncE peptide (F116D) abolishes recruitment to endosomal structures. *Represents the inclusion. Scale bar 20 µm.*

Like wild-type GFP-SNX5 the GFP-SNX5 (F136A) mutant was recruited to punctate endosomal structures throughout the cytoplasm of these cells, and in addition was able to co-immunoprecipate endogenous SNX1 in heterodimeric complexes identically to the wild-type GFP-SNX5 protein (**Fig. 26B**). This implies that BAR-domain mediated heterodimer formation with SNX1 or SNX2 is required for endosomal recruitment, and is not perturbed by the IncE-binding mutation in the PX domain.

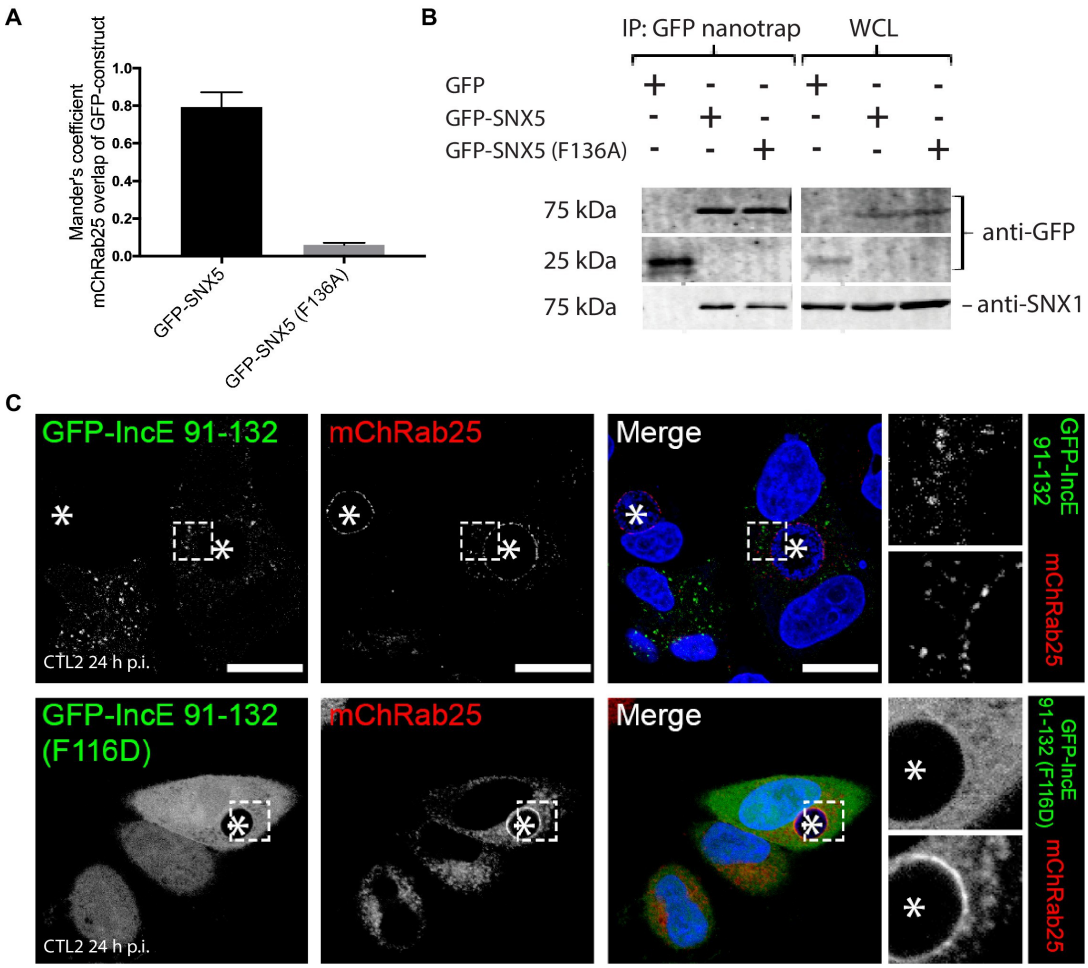


Figure 26. *GFP-IncE C-terminal domain is localised to endosomes but not inclusions. (**A**) Quantitation of the degree of overlap between the GFP-SNX5 constructs and the mCherry-Rab25 inclusion membrane marker from Figure 20A. Mander's correlation coefficient of mChRab25 signal over GFP-SNX5 signals (10 cells per group; error bars, S.D). (**B**) Co-immunoprecipitation of GFP-SNX5 from HeLa cells shows that both the wild-type and mutant protein (F136A) interact equally with endogenous SNX1. This indicates that both proteins are correctly folded and otherwise functional. (**C**) HeLa cells stably expressing mCherry-Rab25 (red) were transfected transiently with GFP-IncE(91-132) or GFP-IncE(91-132)(F116D) (green) and infected with C. trachomatis L2 for 18–24 hr. Mutation in the SNX5 binding IncE peptide (F116D) abolishes recruitment to endosomal structures. Neither construct is recruited to inclusions, which is consistent with the lack of transmembrane regions. *Represents the inclusion. Scale bar 20 µm.*

Lastly, we tested the importance of IncE residues for SNX interaction *in situ* by expressing the GFP-tagged IncE C-terminal domain. The wild-type GFP-IncE(91-132) was recruited to endosomal structures via its interaction with SNX5-related proteins in both uninfected and infected HeLa cells (**Fig. 25B; Fig. 26C**). In contrast however, GFP-IncE(91-132)(F116D), a SNX5-binding mutant, was exclusively cytosolic. Note that neither IncE construct is localised to the inclusion, as expected due to lack of signal peptides and transmembrane domains (**Fig. 26C**).

2.4.4 A model for SNX-BAR recruitment to the inclusion membrane by IncE

Superposition of the SNX5-IncE complex with the SNX5 PX domain in the apo state (25) reveals a significant conformational change in the α-helical extension, as well as localized alterations in the loop between the β1 and β2 strands to accommodate peptide binding (**Fig. 27A**). In essence the IncE β-hairpin acts as a tether between the core PX fold and extended α-helical hairpin, pulling the two sub-structures closer together. Overall the α-helical extension undergoes a maximal movement of ~8-10 Å at the furthest tip, facilitated by the flexibility of the structure following the Pro-rich linker and an apparent hinge-point at Pro97 (**Fig. 27A upper panel**). In the immediate vicinity of Pro97 the SNX5 loop that encompasses Asp43 is significantly shifted and stabilized by the repositioning of Arg103. At both the start of the first α' helix of the extension and the end of the second α'' helix more subtle changes occur in the positions of Met106, Leu128, Tyr132, Leu133 and Phe136. These changes result in formation of the hydrophobic pocket that engages the IncE side-chains Val114, Phe116, Val127 and Leu129 (**Fig. 27A, middle and lower panels**).

To better understand how IncE can recruit the SNX5-containing SNX-BAR complexes to inclusion membranes we constructed an *in silico* model of the SNX5-SNX1 heterodimeric PX-BAR proteins (**Fig. 27B**). Consistent with the length of the IncE C-terminal cytoplasmic sequence the model predicts that the IncE sequence will bind to the surface of SNX5 close to, but oriented away from, the inclusion membrane. While PX domains are commonly able to recognise PtdIns3*P* lipid headgroups, SNX5-related proteins lack the typical binding pocket (**Fig. 27B right panel**), and there is some controversy regarding their ability to mediate specific membrane interactions (8,25). We propose that in the context of *C. trachomatis* infection, SNX5-related proteins are directly associated with the inclusion via IncE protein-protein interactions in a phosphoinositide-independent manner, and are able to recruit their heterodimeric partners SNX1 and SNX2 (87,104,105). The PX-BAR-domain containing complexes are then localised to the

inclusion in a retromer-independent manner (26), and may contribute to the formation of the dynamic inclusion-associated membrane tubules.

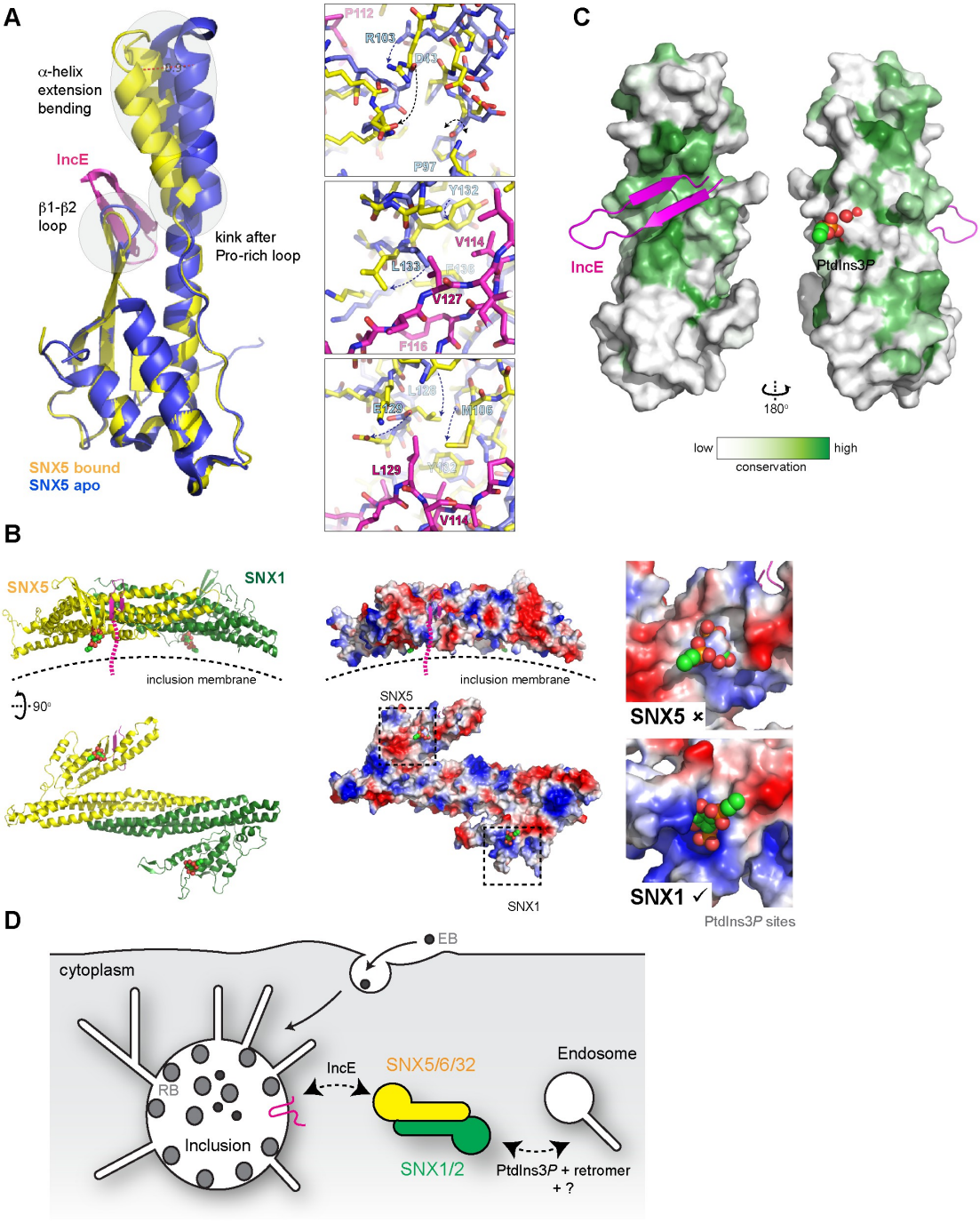


Figure 27. *Conformational changes in SNX5 and a model for SNX-BAR recruitment to inclusion membranes (**A**) Comparison of the SNX5-IncE complex (yellow-magenta) with the previously reported apo- SNX5 PX domain crystal structure (blue) (PDB ID 3HPB)(25). The α-helical extension undergoes a significant displacement in the bound state. The enlarged panels to the right show several close-up views of the binding pocket highlighting conformational changes that are required to accommodate IncE. (**B**) A model for the SNX5-SNX1 PX-BAR heterodimer and its interaction with IncE at the inclusion membrane. The PX-BAR structure was modeled in silico (see*

*methods). The left panel shows cartoon representations of the structure, viewed from the side and from the membrane surface. Middle panels show the same structures in electrostatic surface representation (red, negative; blue positive). The right panels show close ups of the putative PtdIns3P-binding pocket in SNX1 and SNX5, with a PtdIns3P head-group (shown in spheres) docked by aligning the previous SNX9 crystal structure (23). SNX1 has a canonical PtdIns3P pocket, while SNX5 lacks a clear site for lipid head-group binding. (**C**) Sequence conservation of SNX5-related proteins was calculated and plotted using CONSURF. The surface representation indicates exposed side-chains that are evolutionarily conserved in green. The IncE peptide binds to a highly conserved surface groove, while the putative phosphoinositide binding region (25) on the opposite face is neither highly conserved nor poised to allow docking. (**D**) Cartoon model depicting the recruitment of SNX5 and related proteins to the inclusion membrane. Heterodimers with SNX1 or SNX2 will be recruited via IncE in infected cells, and this recruitment will be in competition with the binding of SNX1 and SNX2 to PtdIns3P for normal endosomal association, as well as interactions with other proteins including retromer and unidentified molecules that potentially bind to the conserved groove of the SNX5 PX domain.*

Interestingly, when a cross-species evolutionary analysis of side-chain conservation in the SNX5-related proteins is performed it is clear that the IncE peptide binds a hydrophobic surface groove that is strictly conserved in this protein family (**Fig. 21C**). This very strongly implies that the site is normally engaged in a protein-protein interaction with an as yet unidentified binding partner(s) required for SNX5's regular biological function, and that IncE is directly competing for this interface.

Videos

***Video 1**. Movie showing that wortmannin disrupts SNX5 recruitment to endosomes but not the chlamydial inclusion. HeLa cells stably expressing mCherry-Rab25 (red) were transfected transiently with GFP-SNX5 (green) and infected with Chlamydia trachomatis L2 for 24 hr. Time-lapse videomicroscopy was performed using an interval of 1 min on an inverted Nikon Ti-E deconvolution microscope with environmental control at 40 x magnification. 10 min into recording 100 nM wortmannin was added.*

Video 2. *Animation highlighting the mechanism of interaction between SNX5 and IncE. The SNX5 PX domain is shown in yellow ribbons and the IncE peptide is shown in magenta.*

Table 1. Thermodynamic parameters of IncE binding to SNX PX domains[a]

Sample Cell	Titrant	K_d (μM)	ΔH (kcal/mol)	$T\Delta S$ (kcal/mol)	ΔG (kcal/mol)	N
SNX5 PX	IncE peptide[b]	0.95±0.07	-6.9 ± 0.3	-1.9±0.05	-8.2±0.01	1.01±0.01
SNX6 PX	IncE peptide	1.13±0.08	-5.0± 0.9	-3.0±1	-8.0±0.07	1.01±0.08
SNX32 PX	IncE peptide	1.15±0.07	-6.9± 0.4	-1.3±0.8	-8.2±0.4	1.06±0.005
SNX1 PX	IncE peptide	No binding				

a. Values are the mean from three experiments ±SEM.
b. IncE synthetic peptide sequence PANGPAVQFFKGKNGSADQVILVTQ.

Table 2. ITC data for SNX5 PX domain binding to truncated and mutated IncE peptides[a]

Protein	Peptide	Sequence	*K*d (μM)	Δ*H* (kcal/mol)	TΔ*S* (kcal/mol)	Δ*G* (kcal/mol)	N
SNX5 PX	IncE_1	PANGPAVQFFKGKNGSADQVILVTQ	0.95±0.07	-6.9 ± 0.3	-1.9±0.05	-8.2±0.01	1.01±0.01
	IncE_2	ANGPAVQFFKGKNGSADQVILVTQ	1	-5.0	-2.6	-8.1	0.98
	IncE_3	NGPAVQFFKGKNGSADQVILVTQ	0.93	-6.7	-1.4	-8.1	1.03
	IncE_4	GPAVQFFKGKNGSADQVILVTQ	0.87	-6.8	-1.2	-8.2	1.03
	IncE_5	PAVQFFKGKNGSADQVILVTQ	2	-5.9	-1.2	-8.3	0.99
	IncE_6	AVQFFKGKNGSADQVILVTQ	/	/	/	/	/
	IncE_7	VQFFKGKNGSADQVILVTQ	2.2	-6.9	-1.1	-7.7	0.99
	IncE_8	QFFKGKNGSADQVILVTQ	No binding	/	/	/	/
	IncE_9	FFKGKNGSADQVILVTQ	No binding	/	/	/	/
	IncE_10	FKGKNGSADQVILVTQ	No binding	/	/	/	/
	IncE_11	KGKNGSADQVILVTQ	No binding	/	/	/	/
	IncE_12	GKNGSADQVILVTQ	No binding	/	/	/	/
	IncE_13	KNGSADQVILVTQ	No binding	/	/	/	/
	IncE_14	PANGPAVQFFKGKNGSADQVILVT	0.72	-5.1	-1.6	-8.4	1
	IncE_15	PANGPAVQFFKGKNGSADQVILV	0.97	-6.5	-1.3	-8.2	0.98
	IncE_16	PANGPAVQFFKGKNGSADQVIL	1.1	-5.6	-1.4	-8.12	0.99
	IncE_17	PANGPAVQFFKGKNGSADQVI	8.7	-2.7	-2.5	-6.9	0.99
	IncE_18	PANGPAVQFFKGKNGSADQV	No binding	/	/	/	/
	IncE_19	PANGPAVQFFKGKNGSADQ	No binding	/	/	/	/
	IncE_20	PANGPAVQFFKGKNGSAD	No binding	/	/	/	/
	IncE_21	PANGPAVQFFKGKNGSA	No binding	/	/	/	/
	IncE_22	PANGPAVQFFKGKNGS	No binding	/	/	/	/
	IncE_23	PANGPAVQFFKGKNG	No binding	/	/	/	/
	IncE_24	PANGPAVQFFKGKN	No binding	/	/	/	/
	IncE Q115A	PANGPAV**A**FFKGKNGSADQVILVTQ	6.3	-5.3	-1.6	-6.9	0.90
	IncE F116D	PANGPAVQ**A**FKGKNGSADQVILVTQ	No binding				
	IncE K118A	PANGPAVQFF**A**GKNGSADQVILVTQ	2.8	-6.0	-1.5	-7.5	0.91
	IncE V127D	PANGPAVQFFKGKNGSADQ**D**ILVTQ	No binding				
SNX5 PX L133D	IncE_1	PANGPAVQFFKGKNGSADQVILVTQ	No binding				
SNX5 PX F136A	IncE_1	PANGPAVQFFKGKNGSADQVILVTQ	No binding				
SNX5 PX E144A	IncE_1	PANGPAVQFFKGKNGSADQVILVTQ	15	-9.9	-3.1	-13	0.99

a. Except for IncE_1 all other peptide-binding experiments were performed only once.

Table 3. Summary of crystallographic structure determination statistics[a]

Crystal	SNX5 PX-IncE Form 1	SNX5 PX-IncE Form 2	SNX5 PX-IncE Form 3
PDB ID	5TGI	5TGJ	5TGH
Data collection			
Wavelength (Å)	0.95370	0.95370	0.95370
Space group	$P2_12_12_1$	I2	$P3_2$
Cell dimensions			
a, *b*, *c* (Å)	60.7, 67.5, 88.2	58.4, 80.3, 94.6	100.6, 100.6, 71.7
α, β, γ (°)	90, 90, 90	90, 97.2, 90	90, 90, 120
Resolution (Å)	60.7-1.98 (2.03-1.98)	31.9-2.6 (2.72-2.60)	50.3-2.80 (2.95-2.80)
R_{merge}	0.104 (0.525)	0.153 (0.659)	0.101 (0.713)
R_{meas}	0.112 (0.572)	0.18 (0.777)	0.124 (0.873)
R_{pim}	0.042 (0.225)	0.096 (0.408)	0.051 (0.363)
$\langle I \rangle / \sigma I$	12.4 (3.4)	39.6 (3.2)	11.7 (2.3)
Total number reflections	178868 (11000)	46691 (5757)	115149 (16861)
Total unique reflections	26075 (1805)	13432 (1632)	20001 (2923)
Completeness (%)	100 (100)	99.9 (100.0)	100 (100)
Multiplicity	6.9 (6.1)	3.5 (3.5)	5.8 (5.8)
Half-set correlation (CC(1/2))	0.997 (0.868)	0.986 (0.55)	0.997 (0.683)
Refinement			
Resolution (Å)	45.1- 1.98 (2.02-1.98)	31.9-2.6 (2.69-2.60)	41.2-2.8 (2.87-2.80)
No. reflections/No. R_{free}	26021/ 2000	13421/1342 (1208/134)	19975/1972 (1301/144)
R_{work}/R_{free}	0.192/0.214 (0.221/0.246)	0.199/0.242 (0.276/0.332)	0.236/0.254 (0.329/0.372)
No. atoms			
Protein	2579	2619	5189
Solvent	281	69	0
Average *B*-factor (Å^2)	31.8	42.5	56.0
R.m.s deviations			
Bond lengths (Å)	0.012	0.011	0.015
Bond angles (°)	1.27	1.15	1.27

2.5 Discussion

Although more than fifty putative Incs have been identified in *C. trachomatis*, the exact roles of these inclusion membrane proteins are still poorly understood. *Chlamydiae* manipulate the host cellular and signaling networks via interactions between the cytoplasmic region of Incs and numerous host cell proteins. Recent studies reported retrograde trafficking proteins as significant components of the inclusion, with sorting nexin family members being particularly enriched (26,27). In this study, we present the first reported crystal structure of a chlamydial inclusion protein (IncE) binding to its host effector protein (SNX5). While the detailed mechanism of IncE-mediated protein recruitment will be specific to this family member, the principle of extended cytoplasmic Inc sequences engaging with cellular host proteins on the inclusion is certain to be a general one. A simple analogy would be to consider the Inc proteins as being like a molecular 'velcro' that recognises and attaches host machinery needed for bacterial replication and survival.

The manipulation of endocytic transport machinery is clearly critical for the obligate intracellular survival of *C. trachomatis* (26,27,31). In addition to *C. trachomatis,* SNX1, SNX2, SNX5, SNX6 and the associated retromer complex have also been directly implicated in the cellular pathogenesis of *Coxiella burnetii* (37), *Salmonella* enterica serovar *Typhimurium* (106), hepatitis C virus (107), human papilloma virus (108,109), and *Legionella pneumophila* (110). Broadly then the manipulation of SNX proteins and endosomal trafficking machinery by viral and bacterial pathogens is a common occurrence during intracellular infection, and points to a wide-ranging role in host-pathogen interactions.

Typically PX domains of sorting nexins, including SNX1 and SNX2 (111,112), play an important role in endosomal membrane recruitment by binding the endosome-enriched lipid PtdIns3*P* through four conserved residues (8,38). These residues are conserved in most PX domains including in SNX1 and SNX2, but are entirely absent in SNX5, SNX6 and SNX32. Although there is evidence for the weak association of the SNX5 PX domain with the lipid PtdIns(4,5)P_2 from nuclear magnetic resonance (NMR) spectroscopy experiments (25), the crystal structure does not point to a clear binding mechanism. A second feature that sets SNX5-related proteins apart from the rest of the SNX family is the presence of an extended α-helical insertion. Our work confirms the central importance of this unique insert for the binding of the IncE inclusion protein, and provides the first clear description of how a PX domain can function as a protein-protein interaction scaffold as opposed to a lipid-binding domain.

The high degree of conservation in the IncE binding surface of SNX5 implies that this site is critical for the normal function of SNX5 and its homologs. Previously, the expression of a GFP-tagged IncE C-terminal domain was shown to interfere with the SNX5/SNX6-dependent retrograde trafficking of the cation-independent mannose-6-phosphate receptor (CI-MPR) (26). Combined with our structural data, this infers that IncE is mimicking and interfering with SNX5/SNX6-mediated protein interactions, with a ligand(s) required for normal endosomal trafficking that remains to be discovered. Once recruited to the inclusion, SNX-BAR proteins are localized to the bulk membrane and dynamic tubules. While it is logical to imagine they could play a positive role in the sculpting of the inclusion, this is somewhat difficult to reconcile with the effect of SNX5 and SNX6 knockdown, which results in an increased production of *C. trachomatis* infectious progeny. Alternatively, although a pool of SNX5/SNX6 and associated SNX1/SNX2 proteins remain on endosomes in *C. trachomatis* infected cells, their sequestering by the chlamydial inclusion may interfere with normal endosomal trafficking (**Fig. 27D**). It was thus proposed that the role of IncE could be to compete for SNX-retromer endosomal interactions, resulting in the breakdown of normal trafficking of the CI-MPR and lysosomal hydrolases and hence perturbation of the endolysosomal system's capacity for bacterial destruction (26,27). Defining the precise role of SNX proteins and other endocytic machinery in chlamydial infection will clearly require further study.

In conclusion, our work provides novel molecular insights into the mechanism of SNX protein coercion by the IncE chlamydial effector, and presents a blueprint for future studies of other inclusion protein activities. In addition, our results provide a possible clue to understanding how SNX5-related molecules mediate protein interactions required for canonical cell trafficking pathways.

Data deposition

Structural data are deposited in the protein data bank (PDB) under accession numbers 5TGI, 5TGJ, and 5TGH. Raw diffraction images are available on the University of Queensland eSPACE server (http://espace.library.uq.edu.au/view/UQ:409277).

Author contributions

BP, Conceptualization, Validation, Investigation, Visualization, Methodology, Writing—original draft, Writing—review and editing; HSK, Data curation, Validation, Investigation, Visualization, Writing-original draft, Writing-review and editing; MCK, Data curation, Formal analysis, Supervision, Validation, Investigation, Visualization, Methodology, Writing—

original draft, Writing—review and editing; WMH, Data curation, Validation, Investigation, Writing—original draft, Writing—review and editing; RDT, Conceptualization, Resources, Data curation, Formal analysis, Supervision, Validation, Investigation, Visualization, Methodology, Writing—original draft, Project administration, Writing—review and editing; BMC, Conceptualization, Resources, Formal analysis, Supervision, Funding acquisition, Investigation, Visualization, Writing—original draft, Project administration, Writing—review and editing.

Acknowledgements

The authors would like to acknowledge support from the staff and facilities of the University of Queensland Remote Operation Crystallization and X-ray (UQ ROCX) facility, and the Australian Synchrotron. Microscopy was performed at the Australian Cancer Research Foundation (ACRF)/Institute for Molecular Bioscience (IMB) Dynamic Imaging Facility for Cancer Biology. Elements of this research utilised equipment and support provided by the QLD node of the National Biologics Facility (www.nationalbiologicsfacility.com), an initiative of the Australian Government being conducted as part of the NCRIS National Research Infrastructure for Australia. This work is supported by the Australian Research Council (ARC) (DP0985029; DP150100364) and National Health and Medical Research Council (NHMRC) (APP1058734; 606788). RDT is supported by an NHMRC Senior Research Fellowship (APP1041929), MCK is supported by an Australian Research Council Discovering Early Career Researcher Award (DE120102321), and BMC is supported by an NHMRC Career Development Fellowship (APP1061574).

Chapter 3

Chapter 3

Recombinant expression of the SNX32 PX and SNX6 PX domains and their structural characterisation

3.1 Introduction and Significance

SNX32, a protein primarily expressed in neurons, is a SNX-BAR family member that is highly homologous to SNX5 and SNX6. As mentioned in chapter II, these proteins share a unique structure within their conserved phox homology (PX) domain, which sets them apart from the rest of the SNX protein family. In the previous chapter I reported the three-dimensional X-ray crystallographic structure of the human SNX5 PX domain in complex with the IncE protein from *Chlamydia trachomatis* (113). The PX domain of SNX5 possesses a helix-turn-helix structural insert (25), which is not found in any other SNX family members except for SNX6 and SNX32 (**Fig. 33A**). SNX5, SNX6 and SNX32 proteins were shown to be recruited to Chlamydial infected cells and data from biophysical studies confirmed that a common structure within the SNX5, SNX6 and SNX32 PX domains is required for IncE interaction. The mutation of the key Phe116 side chain in the IncE peptide blocks binding to both the SNX6 and SNX32 PX domains as it does to the interaction with the SNX5 PX domain (113).

In this chapter, using X-ray crystallography, I demonstrate and confirm that the SNX32 PX domain is closely structurally homologous to SNX5, and that it binds to IncE in an essentially identical manner. For this I have successfully performed recombinant expression and purification of the SNX32 PX domain, crystallised it in complex with a peptide derived from IncE and solved its crystal structure at a resolution of 2.3 Å. Although I was able to express and purify the SNX6 PX domain, to date I have only been able to grow microcrystals of this protein in complex with IncE that have so far resisted optimisation.

As the mode of binding of IncE to SNX32 is found to be highly similar to the SNX5 PX-IncE binding mechanism, it suggests that in this case also IncE might be mimicking endogenous cellular proteins that bind to this highly conserved SNX32 binding site. My structural studies thus confirm the biophysical data that was previously published and also provides indirect support for the idea that SNX32 functions in cargo recycling similarly to the well-characterised retromer-associated SNX-BAR proteins SNX6 and SNX6.

3.2 Material and methods

3.2.1 Constructs and Peptides

All synthetic peptides used for protein crystallisation were purchased from Genscript (USA). For the crystallisation set up, peptides were weighed and dissolved in water to make a stock peptide concentration of 10 mM. For crystallisation, this was diluted down to x2 molar excess to the protein molar concentration.

The pGEX-4T-2 bacterial expression plasmids encoding the human SNX32 PX domain (residues 17-166) and the human SNX6 PX domain (residues 29-170) were synthesised and cloned by Genscript (USA). The constructs are N-terminally GST-tagged with site- specific thrombin cleavage sequence located between the GST moiety and the protein of interest.

3.2.2 Transformation of expression plasmids using heat shock method

Plasmid DNA encoding hSNX32PX and hSNX6PX were transformed into *E. coli* BL21 CodonPlus competent cells with ampicillin and chloramphenicol (Amp+/Cm+) antibiotics. The competent cells were allowed to thaw on ice. Transformation was performed by transferring ~2-3 µl of around 0.1-1 µg of the particular cDNA into ~70 µl of thawed competent cells. The tube was gently tapped to mix. This mixture was incubated on ice for 15 min, which was followed by heat shocking of cells at 42°C for 75 s. The incubation was continued on ice for another 2 min. 400 µl of Luria Bertani (LB) media was added into the mixture and was allowed to grow by incubating for 1 h at 37°C. After the incubation, 50 µl of the cells were plated onto the LB agar plates with ampicillin and chloramphenicol antibiotic resistance. The plates were incubated overnight at 37°C.

3.2.3 Recombinant protein expression

A single colony was picked from the LB agar plate and inoculated into LB media with ampicillin (0.1 mg/mL) and chloramphenicol (0.034 mg/mL). The media was then incubated at 37°C with under agitation overnight. The following day, 30 mL from the overnight culture was used to cultivate 1 L LB media containing ampicillin (0.1 mg/mL) and chloramphenicol (0.034 mg/mL) and incubated at 37 °C. Cells were grown until optical density at 600 nm (OD_{600}) reached 0.5-0.6, and expression was induced with 0.5 mM IPTG. The growth is continued at 18°C for 16 h. Cells were harvested by centrifugation of the culture using Beckman rotor JLA 8.1000 (4,000 x g, 15 min, 4°C). The resulting pellets were resuspended in lysis buffer (50 mM Tris pH 8.0, 500 mM NaCl, 5% glycerol, 1 mM DTT, 50 µg/mL benzamidine and 1 µg/mL DNAseI) and stored at –80°C until required or

are directly subjected to cell disruption to continue with protein purification.

3.2.4 Recombinant Protein purification

The thawed cell suspension was lysed using a cell disruptor at 27 kPa at 5ºC. The cellular components of the lysed cells were pelleted using JA25.50 rotor at 15000 RPM for 30 min. Purification was performed using a two-step protocol: affinity chromatography with glutathione sepharose followed by gel filtration using Superdex 75 (16/600) column. The supernatant from the centrifugation step was collected and incubated with the equilibrated glutathione sepharose 4B resin (GE Healthcare) for 1 h at 4°C before eluting the flow through. After the flow through was eluted, the resin was washed with 10 bed volumes of equilibration buffer (50 mM Tris pH 8.0, 100 mM NaCl, 5% glycerol, 1 mM DTT). GST tag was cleaved by adding 100 units of Thrombin protease (Sigma-Aldrich) to the resin slurry and by incubating at room temperature overnight. The cleaved protein was eluted using the equilibration buffer and was subjected to further purification by fast protein liquid chromatography (FPLC) on a Superdex 75 (16/600) column (GE healthcare) which was pre-equilibrated using gel filtration buffer (50 mM TRIS pH 8.0 and 100 mM NaCl). Fractions corresponding to the protein of interest were pooled together and concentrated down (Amicon ultra-15 centrifugal units with a cut off of 10 kDa) to the desired protein concentration for further experiments.

3.2.5 Protein crystallisation, data collection and structure determination

SNX32 and SNX6 PX domains were concentrated to 12 mg/ml and 17 mg/ml respectively for crystallisation. These proteins were directly mixed together with the appropriate IncE peptide at a 2:1 molar ratio of peptide to protein and were incubated on ice for 1 h. To increase the chances of identifying a crystallisation condition for crystal nucleation, eight 96-well crystallisation hanging-drop screens were set up using a Mosquito Liquid Handling robot (TTP LabTech) at 20°C in the UQ ROCX facility. After setting up these 96-well screens, the trays were incubated at 20°C in a Rockimager storage hotel (Formulatrix), where the drops were imaged at different time points for 21 days. For SNX32PX, optimized diffraction-quality crystals were obtained using streak seeding in sitting drop vapor diffusion plates. The crystallisation condition for SNX32 PX was 20% PEG 8000, 0.1 M Tris base pH 8.0, 0.01 M $MgCl_2$. Data were collected at the Australian Synchrotron MX1 and MX2 Beamlines, integrated with iMOSFLM (93), and scaled with AIMLESS (94) in the CCP4 suite (95). The structures were initially solved by molecular replacement with PHASER (96) using the SNX5PX-IncE crystal structure as the input model (PDB code 5TGI). The resulting model was rebuilt with COOT (114), followed by repeated rounds of

refinement with PHENIX (98). All structural figures were generated using PyMOL (DeLano scientific).

For SNX6PX domain in complex with IncE, the reservoir condition that produced positive crystal hits in the initial 96-well plate was further optimized in a pH versus precipitant concentration grid screen around the reservoir condition 20% PEG 6000, 0.1 M NaAc pH 5.0, 0.01 M $ZnCl_2$ using a sitting drop vapor diffusion method in a 24-well plate.

3.3 Results

3.3.1 Purification of the SNX6 PX domain

In order to confirm the structural mechanism of SNX6 PX-IncE interaction I successfully expressed and purified SNX6 PX domain as described under 3.2.3 and 3.2.4. Samples from every step of purification were analysed on SDS-PAGE to ensure the presence of the protein of interest.

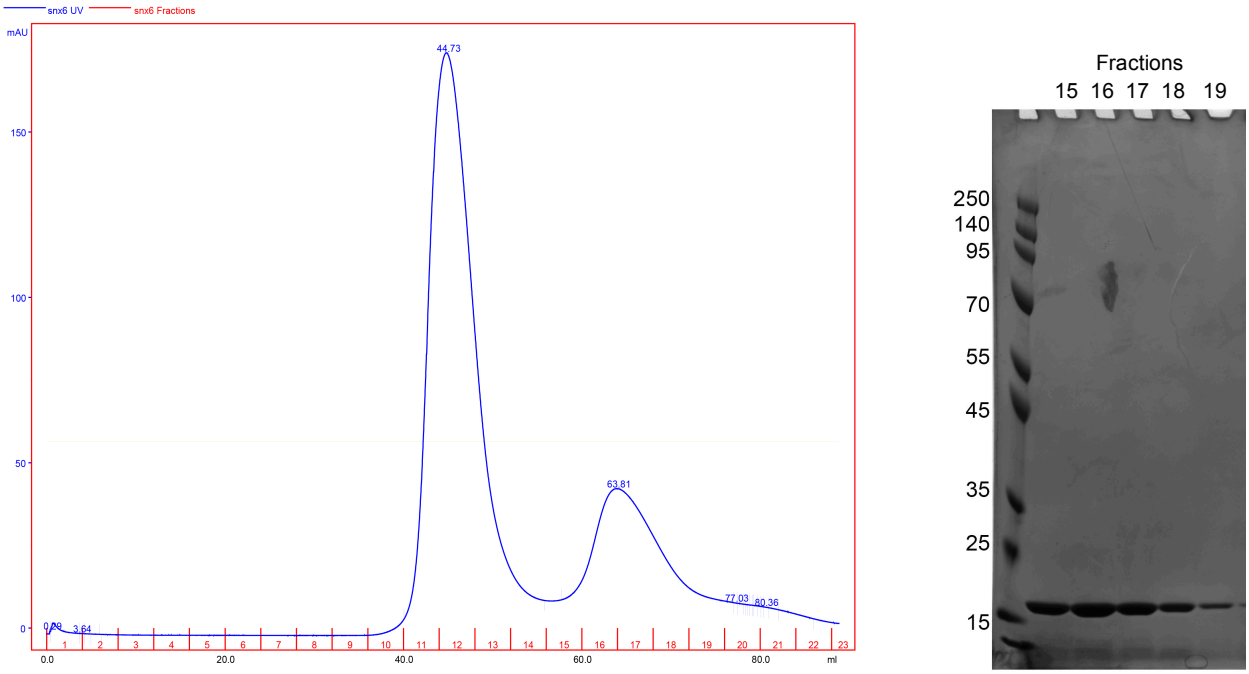


Figure 28. *FPLC chromatography profile of untagged SNX6-PX. The peak between 60-70 mls corresponds to the single bands (MW=16.2 kDa) on the gel (fractions 15-19).*

The theoretical molecular weight of untagged hSNX6-PX is around 16.2 kDa and was expected to elute as a monomeric species (when compared to the column standards) between 60 mL and 70 mL on a Superdex 75 (16/600) column. The affinity purified SNX6 PX was eluted as one peak at the expected molecular weight. Fractions 15 to 19

corresponding to the elution peak contained relatively pure SNX6 PX as shown in SDS-PAGE analysis (**Fig. 28**). The protein was concentrated to 20 mg/ml for further ITC experiments as described in chapter II, and crystallisation experiments.

3.3.2 Crystallisation trials of SNX6 PX in complex with IncE

Out of eight different initial screens that I tried in 96-well trays, two conditions produced microcrystals of SNX6PX-IncE complex in 2 days. The reservoir conditions were 20% PEG 6000, 0.1 M NaAc pH 5.0, 0.01 M $ZnCl_2$, and 20% PEG 4000, 5 mM $CdCl_2$, and 0.1 M Tris pH 8.0. These conditions were optimised in 24-well plates using a grid screen strategy. While I was able to successfully reproduce the crystals at a larger scale in both the conditions, unfortunately the crystals were still small, highly nucleated or clustered.

A **B** **C**

Figure 29. *Crystals of SNX6PX-IncE complex from 20% PEG 6000, 0.1 M NaAc pH 5.0, 0.01 M $ZnCl_2$ (**A**) Crystals obtained in the 96-well plate. (**B**) and (**C**) Crystals obtained after optimizing in 24-well plate.*

Despite trying several optimisation strategies including decreasing the protein concentrations and precipitant concentrations, streak seeding into diluted reservoir solutions, adding glycerol to slow down the rate of nucleation upon mixing of protein with the conditions, I was not able to produce macrocrystals good enough for a high resolution structural analysis. The crystals from 20% PEG 6000, 0.1 M NaAc pH 5.0, 0.01 M $ZnCl_2$ were too clustered and small to collect any diffraction data (**Fig. 29**). Using the MX2 high intensity microfocus beamline at the Australian Synchrotron I was able to see protein diffraction to a resolution of ~6 Å from microcrystals obtained from the condition 20% PEG 4000, 5 mM $CdCl_2$, and 0.1 M Tris pH 8.0 (**Fig. 30**). While this is promising, the crystals still require further optimisation to yield macrocrystals.

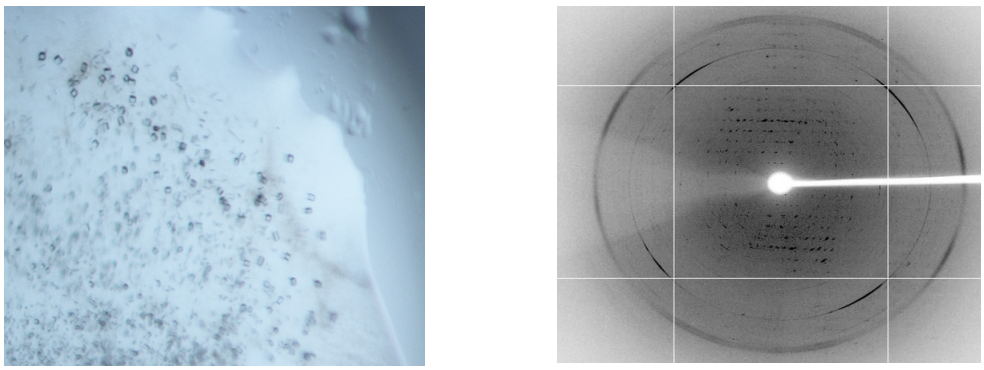


Figure 30. *Crystals of SNX6PX-IncE complex from 20% PEG 4000, 5 mM $CdCl_2$, and 0.1 M Tris pH 8.0 (left panel). Right panel shows that these crystals diffracted to 6 Å (images obtained at the Australian Synchrotron MX2 beamline).*

3.3.3 Purification of SNX32PX domain

As described in Chapter 2 the IncE protein binds equally well to the three proteins SNX5, SNX6 and the poorly characterized brain specific homologue SNX32. **Fig. 31** shows the purification of the SNX32 PX domain after removal of the GST tag using a Superdex 75 column, which yielded recombinant SNX32PX protein with a final purity of more than 95%. The SNX32 PX purification steps are described in greater detail in the materials and methods.

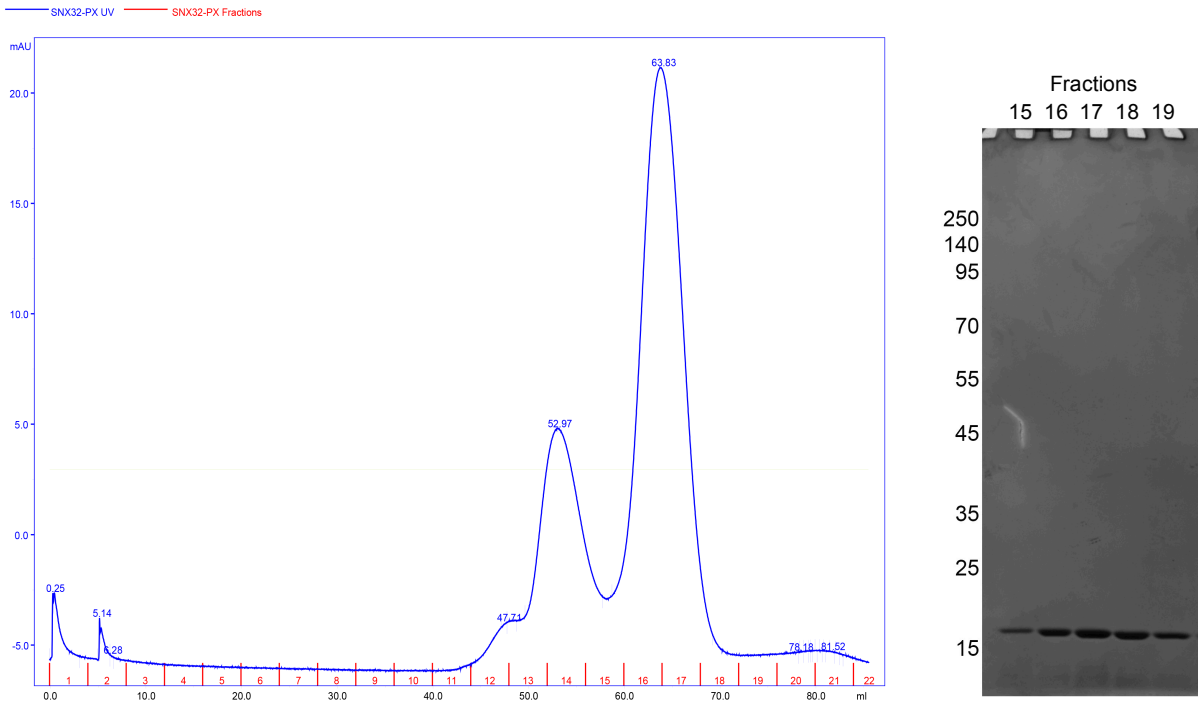
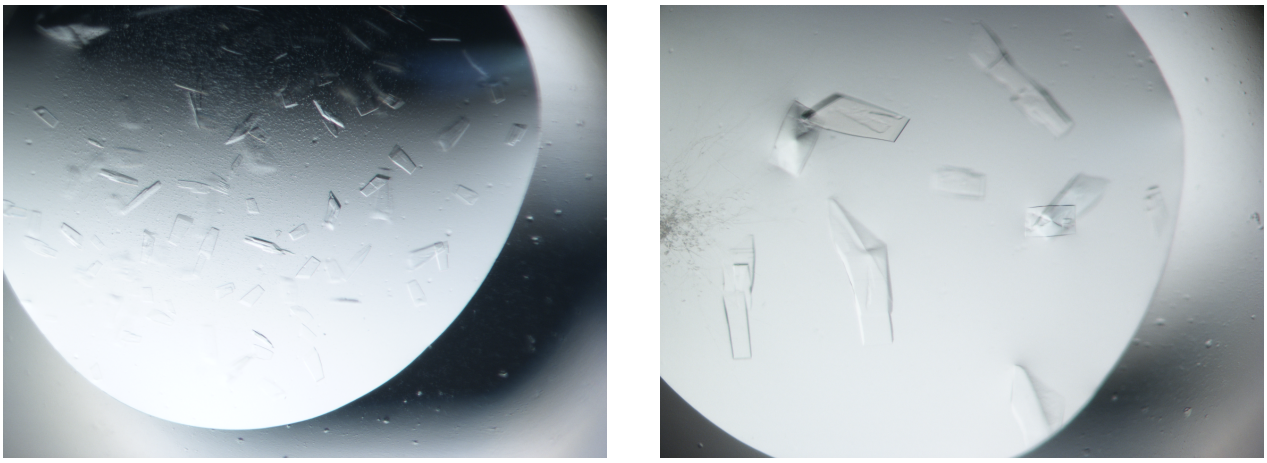


Figure 31. *FPLC chromatography profile of untagged SNX32-PX domain. The fractions (15 to 19) from the major peak around 60-70 mls correspond to the single bands on the gel (MW=15.6 kDa).*

3.3.4 Crystallisation and structural determination of SNX32-PX in complex with IncE

Performing crystallisation screens of the SNX32 PX domain in the presence of the IncE peptide (similar to SNX6 described above) yielded rectangular-shaped and needle-shaped crystals in two conditions out of 760 initial screen conditions using commercial kits.

After trying several optimization strategies including decreasing the protein concentrations and precipitant concentrations, and streak seeding into diluted reservoir solutions, I obtained single macro crystals that were used to obtain X-ray diffraction at the Australian Synchrotron beamlines. The reservoir condition that yielded the diffracting crystal was 20% PEG 8000, 0.1M Tris base pH 8.0, 0.01M $MgCl_2$ (**Fig. 32**).



A **B**

Figure 32. *Crystals of SNX32PX-IncE complex in the reservoir condition 20% PEG 8000, 0.1M Tris base pH 8.0, 0.01M $MgCl_2$ before (**A**) and after (**B**) streak seeding*

After collecting high-resolution diffraction data I then determined the structure of the SNX32-IncE complex using molecular replacement, with the SNX5 PX domain as my initial template structure. The final refined structure of the SNX32 PX domain is highly similar to SNX5 as expected (**Fig. 33B**). It is composed of a three-stranded β-sheet (β1, β2 and β3) followed by three closely packed α-helices. The first and second α-helices are connected by an extended proline-rich sequence. Typically PX domains have been found to bind to the endosome-enriched lipid phosphatidylinositol-3-phosphate (PtdIns3*P*) via a basic pocket formed at the junction between the β3 strand, α1 helix and ΨPxxPxK loop (Pro-rich loop). In contrast SNX5, SNX6 and SNX32 possess major alterations in the PtdIns3P-binding pocket that preclude canonical lipid head-group docking. In addition they possess a unique extended helix-turn-helix insert between the ΨPxxPxK loop and α2 helix of unknown function.

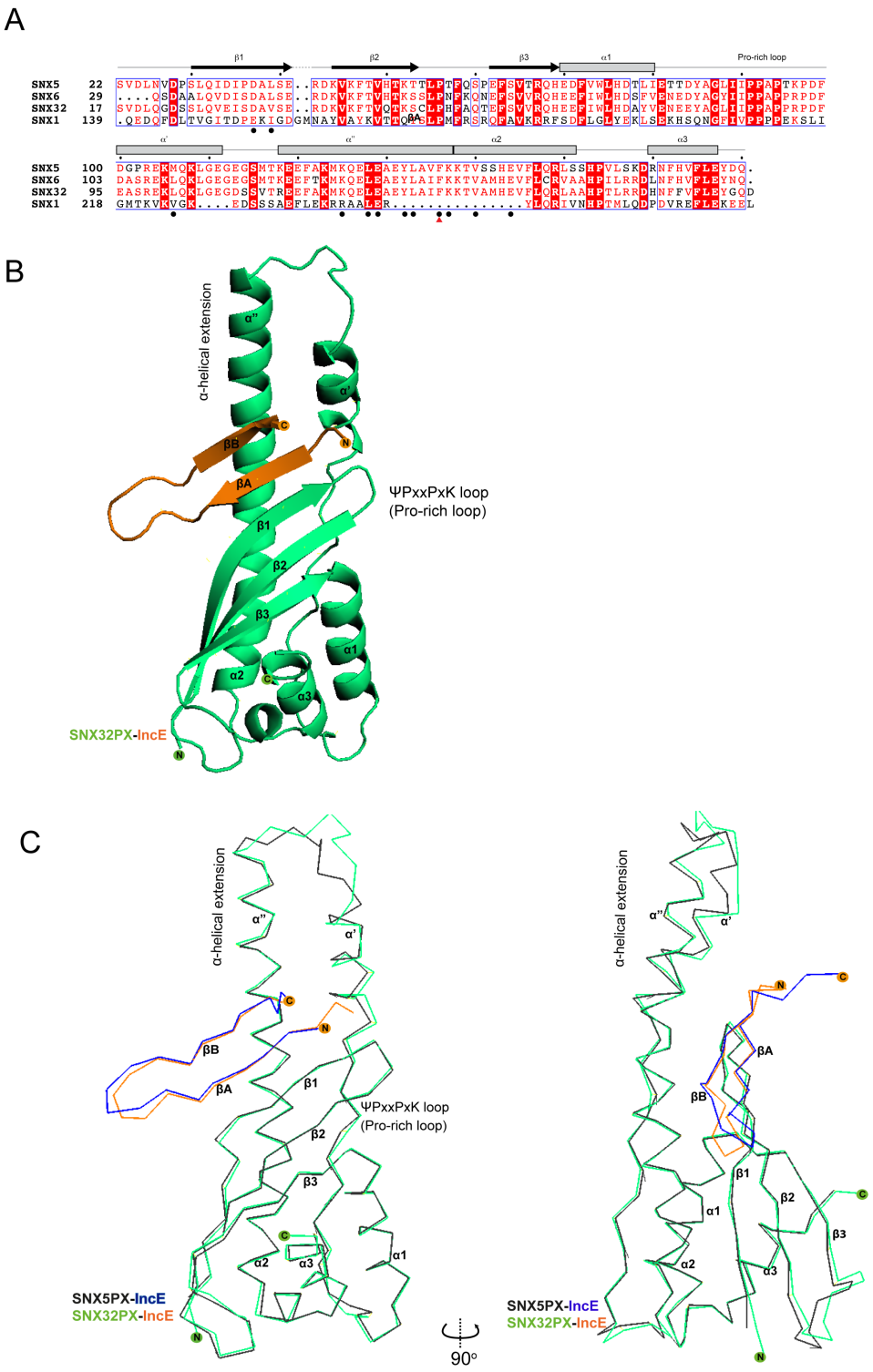


Figure 33. *(**A**) Sequence alignment of human SNX1, SNX5, SNX6 and SNX32 PX domains. Conserved residues are indicated in red. Side-chains that directly interact with IncE in the crystal structure are indicated with black circles. Mutation that block IncE binding is highlighted with red triangle. (**B**) Crystal structure of the SNX32 PX domain (green) in complex with IncE residues 107-132 (orange) shown in cartoon diagram. All structure images were generated using PyMOL (Delano Scientific) **(C)** Superposition of crystal structure of SNX32PX-IncE backbone with SNX5PX-IncE backbone.*

The crystals of SNX32PX-IncE diffracted to a resolution of 2.26 Å and belong to spacegroup I121 (**Table 4**). The superposition of SNX32PX-IncE and SNX5PX-IncE crystal structures reveal the mode of IncE binding in both cases as nearly identical (**Fig. 33C**). Just as in SNX5PX-IncE crystal structure, the IncE sequence forms a long β-hairpin structure that binds within a complementary hydrophobic groove, at the base of the extended α-helical insertion of the SNX32 PX domain and adjacent to the β-sheet sub-domain (**Fig. 33C**). The β-hairpin structure of IncE is directly incorporated as a β-sheet augmentation of the β1, β2 and β3 strands of SNX32. Aside from main-chain hydrogen bonding to form the extended β-sheet, IncE Phe116 forms a π-stacking with the SNX32 Phe136 side-chain similar to what is seen in the SNX5 complex (**Fig. 34A and 34B**). In our previous paper it has been shown that mutation of the IncE Phe116 completely abolishes the binding of SNX32 to the Chlamydial protein IncE (113).

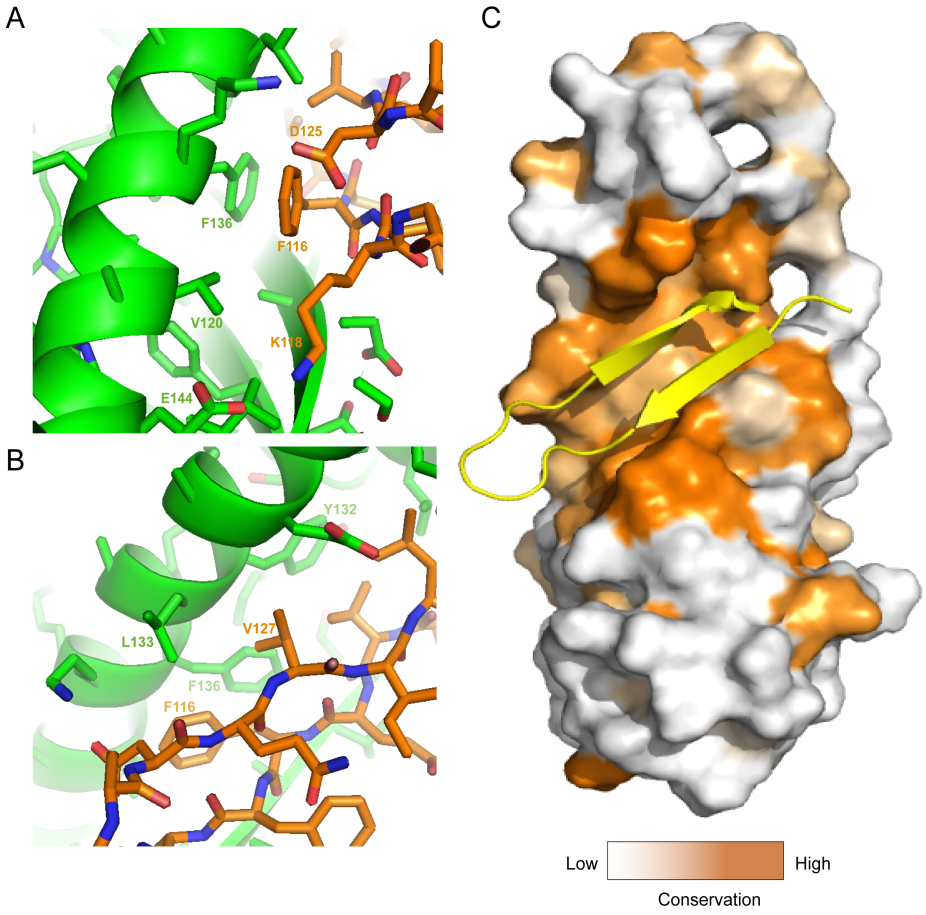


Figure 34. *Interactions between SNX32 (green) and IncE (orange).* ***(A)*** *Close up of the SNX32PX-IncE interface highlighting the important interactions between SNX32PX and the N-terminus of IncE.* ***(B)*** *Close up of the SNX32PX-IncE interface highlighting the important interactions between SNX32PX and the C-terminus of IncE.* ***(C)*** *Sequence conservation of SNX32PX and its related proteins (calculated using CONSURF).*

Table 4. Summary of crystallographic structure determination statistics

Data collection	
Wavelength (Å)	0.95370
Space group	I_{121}
Cell dimensions	
a, *b*, *c* (Å)	90.8, 46.0, 134.2
α, β, γ (°)	90, 105.58, 90
Resolution (Å)	83.68-2.26 (3.27-2.27)
R_{merge}	0.151 (0.894)
R_{meas}	0.176 (1.046)
R_{pim}	0.063 (0.540)
<*I*> / s*I*	10.7 (2.5)
Total number reflections	187434 (15896)
Total unique reflections	25168 (2222)
Completeness (%)	99.7 (97)
Multiplicity	7.5 (7.2)
Half-set correlation (CC(1/2))	0.996 (0.726)
Refinement	
Resolution (Å)	43.77- 2.27 (2.35-2.27)
No. reflections/No. R_{free}	25103/ 1280
R_{work}/R_{free}	0.186/0.235(0.193/ 0.231)
No. atoms	
Protein	2676
Solvent	57
Average *B*-factor (Å^2)	26.3
R.m.s deviations	
Bond lengths (Å)	0.014
Bond angles (°)	1.264

3.4 Discussion

The main objective of this chapter was to understand how the chlamydial inclusion protein IncE binds to SNX32PX and also to compare the binding mechanism with that of SNX5PX-IncE. The binding affinity of the IncE C-terminal sequence (PANGPAVQFFKGKNGSADQVILVTQ) towards SNX32 was found to be K_d ~1.15 μM in chapter 2, which was very similar to the binding affinity of IncE to SNX5 (K_d ~1 μM). This

data basically confirmed that IncE would be binding within the same site. The X-ray crystal structure clearly confirms this model.

In vivo, SNX32 appears to be exclusively expressed in neuronal cells based on microarray studies (biogps.org). Therefore it is unlikely it would ever actually encounter chlamydial pathogens during their infection cycle. However, the fact that SNX32 binds to IncE in the same way as SNX5 leads us to the same conclusion that IncE is mimicking endogenous SNX5, SNX6 and SNX32 binding proteins, which would normally bind to the same conserved groove for their normal function in the cell (**Fig. 34C**). There have been recent reports that suggest the SNX5 may directly interact with transmembrane cargos sorted by the SNX-BAR protein at endosomes via its PX domain. For example, both CIMPR and IGF1R were found to be interacting with the SNX5 PX domain via their cytoplasmic domains (115,116). As described in the following chapter, I now show that these cargos do indeed bind to the same site as IncE, and begin to dissect this normal functional role of the conserved SNX5/SNX6/SNX32 proteins.

Chapter 4

Chapter 4

Structural basis for the interaction of transmembrane cargoes with the SNX-BAR proteins

4.1 Introduction and significance

The retromer complex is a highly conserved multiprotein complex that co-ordinates recycling of various transmembrane cargo proteins from endosomes either to trans-Golgi network (TGN) or back to the plasma membrane. In yeast, retromer is composed of two subcomplexes; i) the vacuolar protein sorting (Vps) core consists of three proteins Vps35, Vps29 and Vps26 and ii) two members of the sorting nexin (SNX) protein family Vps5p and Vps17p. Vps5 and Vps17 contain C-terminal bin-Amphiphysin-Rvs (BAR) domains that form Vps5p/Vps17p heterodimers, and assemble into higher order helical arrays promoting membrane curvature to encourage endosomal membrane tubulation (117). Both subcomplexes contribute towards separate roles during the recycling of cargoes. According to the current working model of retromer-dependent cargo recycling, the Vps core directly binds to the intracellular cytosolic domains of the transmembrane cargo molecule and promotes the enrichment of cargo molecules in endosomal retrieval subdomains, away from the subdomains assigned for protein degradation. Once the local concentration of the selected cargo is high enough, cargoes are coupled by retromer to the Vps5p/Vps17p membrane tubules to form cargo-containing tubular transport carriers, which then enable the recycling of cargoes to different destinations (16,17,118).

In mammalian cells, the core retromer components are very highly conserved and the model of retromer activity is considered to be generally the same as in yeast. The major difference is that the SNX-BAR dimers present in mammalian cells are composed of either SNX1 or SNX2 (orthologues of Vps5p) heterodimers with SNX5 or SNX6 (orthologues of Vps17p) (16,18,87). However, the biochemical evidence for a stable interaction between the Vps core and SNX-BAR sub complexes in mammalian cells has been more elusive than with the yeast counterparts. So far, the accepted connection between these two sub complexes was through their proposed mutual functional role in the retrograde trafficking of cation-independent mannose 6-phosphate (CI-MPR) (119,120). This receptor delivers lysosomal hydrolases from the Golgi to the endosomes, and then is recycled back to the Golgi from endosomes (14,121,122) through a process that was believed to require direct interaction of the CI-MPR tail with the VPS trimer of the

retromer complex (119,120).

This previously accepted model of retromer activity and the co-operation of Vps core and SNX-BAR dimer subcomplexes was recently questioned by two separate studies from the Steinberg and Cullen laboratories (115,116). These two groups found that the Vps35/Vps26/Vps29 core complex was less important for CI-MPR trafficking (and trafficking of a variety of other receptors) than the SNX-BAR dimer heterodimers. Using several knockdown, knockout and rescue experiments, these reports showed that the loss of SNX-BAR dimers results in steady state redistribution of CI-MPR from the Golgi to the endosomes, due to an inability to enter transport tubules. In contrast the loss of Vps trimers did not cause CI-MPR redistribution on endosomes. Simonetti *et al*. and Kvainickas *et al*. also observed in WT cells that the CI-MPR receptor predominantly overlaps with SNX-BAR dimer in an endosomal subdomain spatially separate from Vps trimer, thus contradicting previous studies that reported interaction of CI-MPR with Vps35 (119).

While this work and the relative contributions of the Vps timeric core and the SNX-BAR proteins are still debated (123), it is clear that the SNX-BAR proteins play a key role in cargo recruitment in mammalian cells. The proteomics studies identified a number of important transmembrane cargo proteins whose trafficking is specifically dependent on SNX5 and SNX6 (115,116). In addition to CI-MPR, these cargoes included Insulin-like growth factor 1 receptor (IGF1R) and Semaphorin- 4C precursor (SEMA4C). In chapter 2 and 3, I have reported that the Chlamydial effector protein IncE binds to a conserved hydrophobic groove on the PX domain of SNX5 (113). Elwell *et al*. further showed IncE displaces CI-MPR from SNX5 (124), which leads to the speculation that the same binding site on SNX5 (and possibly SNX6 and the brain specific homologue SNX32) mediates interactions with the transmembrane cargoes. Mutation of the SNX5 PX domain residues F136D and Y132D revealed loss of CI-MPR and IGF1R binding supporting this model (115,116). The WLM motif in the cytoplasmic tail of CI-MPR was found to be required for their binding and sorting by the SNX-BAR dimer (115,120). Since not all cargoes contain this motif, different binding motifs on the cargoes might explain the cargo selectivity for these SNX-BAR proteins.

In unpublished work, our collaborators Dr. Boris Simonetti and Prof. Peter Cullen (Bristol, UK) have mapped the binding of SNX5 to several selected cargoes (CI-MPR, IGF1R and SEMA4C) through immunoprecipitation experiments of various receptor truncations. This work has identified a region with a conserved sequence on all three cargoes V-x-[F/Y]-x-Y-S (where x can be any amino acid) (**Fig. 35**). Our hypothesis is this sequence represents a specific sorting motif mediating direct interaction between the

transmembrane cargoes and the PX domains of SNX5, SNX6 and SNX32.

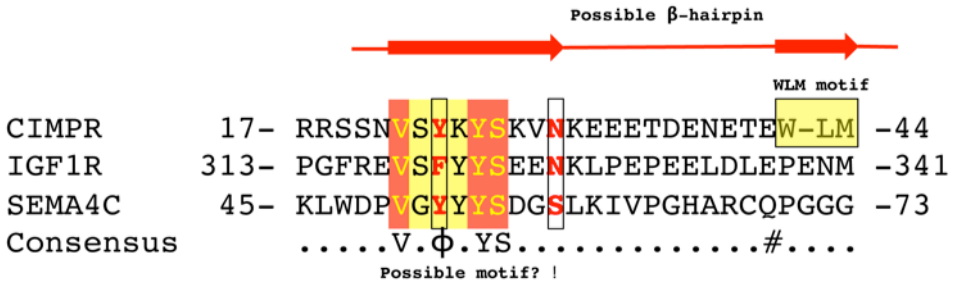


Figure 35. *Sequence alignment of the cytoplasmic tail of CIMPR, IGF1R and SEMA4C showing the conserved V-x-F/Y-x-Y-S sequences that could potentially be the binding site on the transmembrane cargoes.*

In this chapter, using isothermal titration calorimetry (ITC) and X-ray crystallography, I demonstrate and confirm that the PX domains of SNX5, SNX6 and SNX32 directly bind to the conserved cytoplasmic motifs from each of these receptors CIMPR, IGF1R and SEMA4C. I also report the first crystal structure of the PX domain of SNX5 bound to a transmembrane cargo protein, providing the very first structural information on how SNX-BAR proteins can directly regulate transmembrane protein recycling. For this I have successfully performed recombinant expression and purification of the SNX5, SNX6 and SNX32 PX domains and studied their interaction with the cargo peptides using ITC, crystallised the SNX5 PX domain in fusion with CIMPR and solved its crystal structure at a resolution of 3.1 Å.

The results from this chapter validate and confirm the hypothesis from chapter 2 (113), chapter 3 and the recent publications (115,116) that the SNX-BAR proteins bind directly to the transmembrane cargoes, independent of the Vps core subcomplexes. This confirms the that SNX5 PX domain functions as a site for the normal transporting of proteins/cargoes that is fundamental to all living cells. My structural studies will now be used to design specific point mutations in SNX5 and the cargo receptors to confirm the importance of these interactions in receptor recycling in cell assays to be performed in Bristol. Further optimisation of SNX5-CIMPR crystal data and future crystal structures of SNX5-IGF1R fusion and SNX5-SEMA4C complexes will provide molecular insights into the biologically critical endosome to trans-Golgi network (TGN) and endosome-to-plasma membrane trafficking of potentially many different transmembrane cargoes via the SNX-BAR-mediated pathway.

4.2 Material and methods

4.2.1 Peptides and expression constructs

All synthetic peptides used for ITC experiments (**Table. 5**) were purchased from Genscript (USA). For the ITC experiments, peptides were weighed and dissolved in 50 mM HEPES (pH 8.0) and 100 mM NaCl (ITC buffer) to make a stock peptide concentration of 2 mM, which was diluted to 0.75 mM before use.

All bacterial expression constructs for human SNX proteins were synthesized and cloned into pGEX-4T-2 by Genscript (USA). These included the SNX5PX-CIMPR fusion (SNX5 residues 26–170 with CI-MPR residues 17–44 fused at the C-terminus), SNX5PX-IGF1R fusion (SNX5 residues 26–170 with IGF1R residues 313-341 fused at the C-terminus) SNX5PX-SEMA4C fusion (SNX5 residues 26–170 with SEMA4C residues 45-73 fused at the C-terminus), SNX32PX-CIMPR (SNX32 residues 17-166 with CIMPR residues 17-44 fused at C-terminus) and SNX5PX (residues 26-170) F136D mutant. The constructs are N-terminally GST-tagged with site- specific thrombin cleavage sequence located between the GST moiety and the protein of interest.

Table 5. List of transmembrane cargo peptides used in this chapter

Peptides	Sequences
CIMPR_1	SNVSYKYSKVNKEEETDENETEWLMEEIQ
CIMPR_2	SNVS**N**KYSKVNKEEETDENETEWLMEEIQ
CIMPR_3	SNVSYKYSKVNKEEETDENETE**AAA**EEIQ
IGF1R_1	GFREVSFYYSEENKL
IGF1R_2	RE**A**S**A**Y**A**SEENKLPE
SEMA4C	WDPVGYYYSDGSLKIVPGHARCQPG

4.2.2 Recombinant protein expression and purification

All constructs were expressed in *Escherichia coli* BL21 Codon Plus supplemented with appropriate antibiotics. Single colonies from cultures grown on LB agar plates were inoculated into 50 mL Luria-Bertani (LB) broth with ampicillin (0.1 mg/mL) and chloramphenicol (0.1 mg/mL), and grown at 37°C with shaking overnight. The following day, 30 mL from the overnight culture was used to inoculate 1 L LB media containing ampicillin (0.1 mg/mL) and chloramphenicol (0.1 mg/mL) and incubated at 37°C. Cells were grown to an optical density (OD) of 0.7-0.8 at 600 nm and induced with 0.5 mM isopropyl-β-D-thiogalactopyranoside (IPTG). Cultures were incubated with shaking

overnight at 18°C. Cells were harvested using a Beckman rotor JLA 8.1000 at 4000 RPM for 15 min at 4°C. Pellets were resuspended in 10 mL lysis buffer (50 mM Tris (pH 8.0), 100 mM NaCl, 5% glycerol, 1 mM DTT, 0.1 mg/ml benzamidine, 0.1 mg/ml DNase) per litre of culture. The cells were subjected to cell disruption and centrifugation using JA 25.50 at 18,000 RPM for 30 min at 4°C. The soluble fractions were first purified using affinity chromatography with glutathione-sepharose, and the GST tags were cleaved by thrombin while still bound to the column. The proteins were eluted in 50 mM Tris (pH 8.0), 100 mM NaCl, 5% glycerol, and 1 mM DTT, and then further polished using gel filtration chromatography (Superdex 75 or 200, GE healthcare) in a buffer containing 50 mM HEPES (pH 8.0), 100 mM NaCl. The fractions corresponding to the respective proteins were then pooled and concentrated for ITC experiments or crystallization.

4.2.3 Isothermal titration calorimetry (ITC)

ITC experiments were performed on a MicroCal PEAQ-ITC instrument at 25°C. The proteins were buffer exchanged into ITC buffer (50 mM HEPES (pH 8.0) and 100 mM NaCl) by gel filtration prior to ITC experiments. Transmembrane cargo peptides at 0.75 mM were titrated into 50 mM PX domain samples. The binding data was processed using MicroCal PEAQ-ITC analysis software with a single site-binding model to determine the stoichiometry (n), the binding constant (K_d), the estimated heat of binding (ΔH), the Gibbs free energy (ΔG) and the binding entropy (ΔS). Three experiments were performed for each set of samples to determine the average ± standard error of the mean (SEM) for thermodynamic quantities, except for the peptide mutations and SNX32PX interaction experiments where only single experiments were performed. For these single experiments, all experiments were performed using the same batch of protein to allow direct comparisons to be made.

4.2.4 Protein crystallization, data collection and structure determination

All proteins were concentrated to 12 mg/ml for crystallization. Four 96-well crystallization hanging-drop screens were set up using a Mosquito Liquid Handling robot (TTP LabTech) at 20°C in the UQ ROCX facility. After setting up these 96-well screens, the trays were incubated at 20°C in a Rockimager storage hotel (Formulatrix), where the drops were imaged at different time points for 21 days. Data were collected at the Australian Synchrotron MX2 Beamline, integrated with iMOSFLM (93), and scaled with AIMLESS (94) in the CCP4 suite (95). The structures were initially solved by molecular replacement with PHASER (96) using the SNX5PX-IncE crystal structure as the input model (PDB code 5TGI) (113). The resulting model was rebuilt with COOT (114), followed by repeated

rounds of refinement with PHENIX (98). All structural figures were generated using PyMOL (DeLano scientific).

4.3 Results

4.3.1 Purification of SNX5PXF136D mutant construct

To understand the interaction between SNX5PX mutant and different transmembrane cargoes through ITC, I successfully expressed and purified SNX5PXF136D mutant construct to be used as a negative control. **Fig. 36** shows the purification of this construct after the removal of the GST tag using a Superdex 200 column. The theoretical molecular weight of untagged SNX5PXF136D is around 17 kDa and was eluted as monomeric species (when compared to the column standards) between 80 mL and 100 mL on a Superdex 200 (16/600) column.

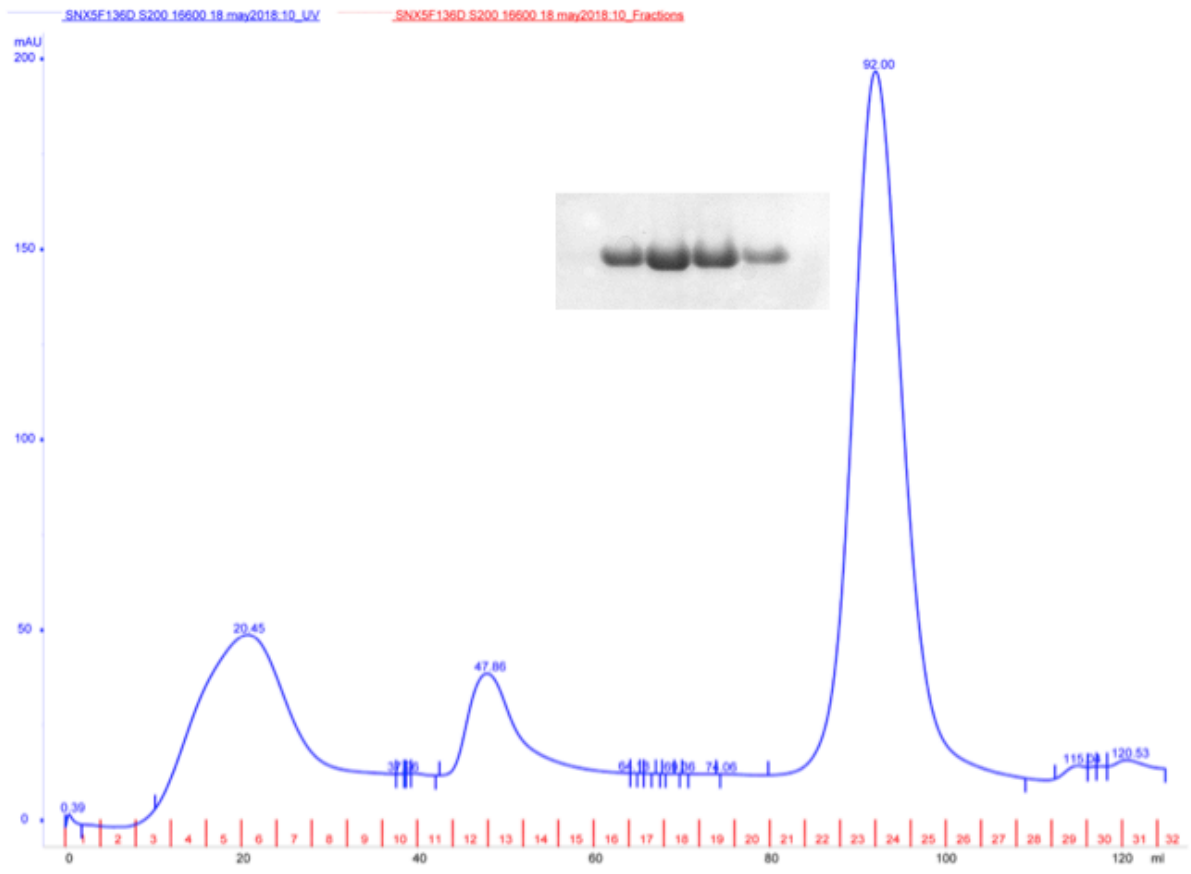


Figure 36. *FPLC chromatography profile of untagged SNX5PXF136D mutant. The fractions from 22 to 25 from the major peak around 80-100 mls correspond to the single bands on the gel (MW=17 kDa).*

4.3.2 Interaction studies between the PX domains of SNX5, SNX6, SNX32 and SNX5PXF136D with the cytoplasmic tail of transmembrane cargoes

Having observed interaction of PX domains of SNX5, SNX6 and SNX32 and transmembrane cargoes through immunoprecipitations (Simonetti and Cullen, *personal communication*), I investigated whether these SNX-BAR proteins can interact directly with the selected transmembrane cargoes CI-MPR_1, IGF1R_1 and SEMA4C using ITC (**Fig. 37**).

The experiments were set up to describe binding of synthetic peptides mimicking cytoplasmic tails of transmembrane cargoes to the recombinantly produced PX domains of SNX-BAR proteins using isothermal titration calorimetry (ITC). The details of synthetic peptides are presented in **table 5** under 4.2.1.

PX domains of SNX5, SNX6 and SNX32 bound directly to CI-MPR_1 with similar affinities, with K_d ≈ 26 µM, 25 µM and 28 µM respectively (**Fig. 37A**). While SNX5PX and SNX32PX bound to IGF1R_1 (K_d ≈ 16 µM and 19 µM respectively), SNX6PX failed to do so (**Fig. 37B**). Interestingly, this also correlates with an inability of SNX6 to co-immunoprecipitate with the IGF1R in cell lysates (Simonetti and Cullen, *personal communication*). SNX5, SNX6 and SNX32 (K_d ≈ 14 µM, 13 µM and 12 µM respectively) bound to SEMA4C with similar affinities (**Fig. 37C**).

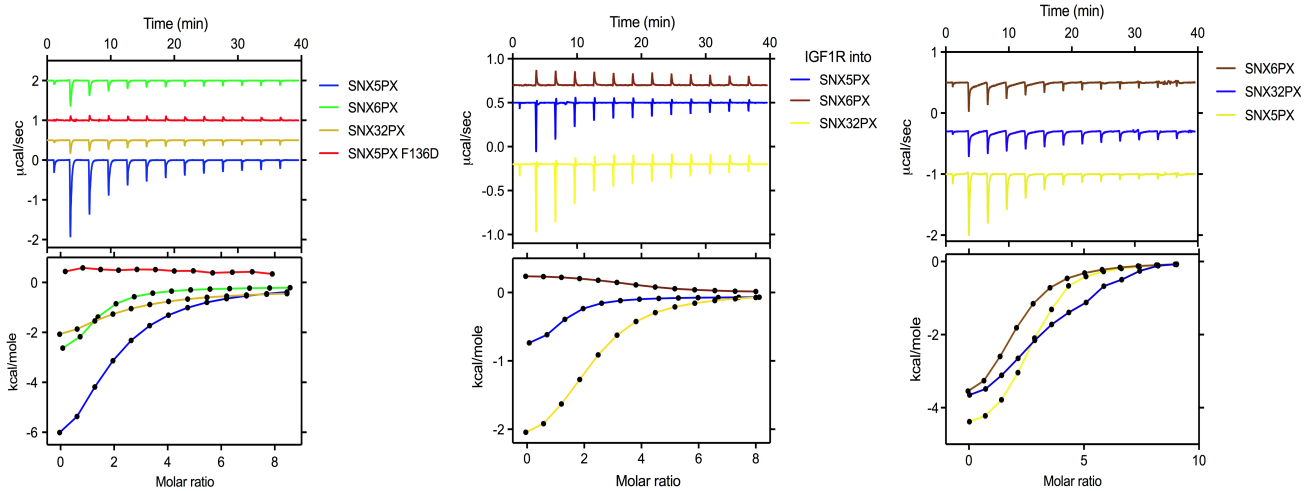
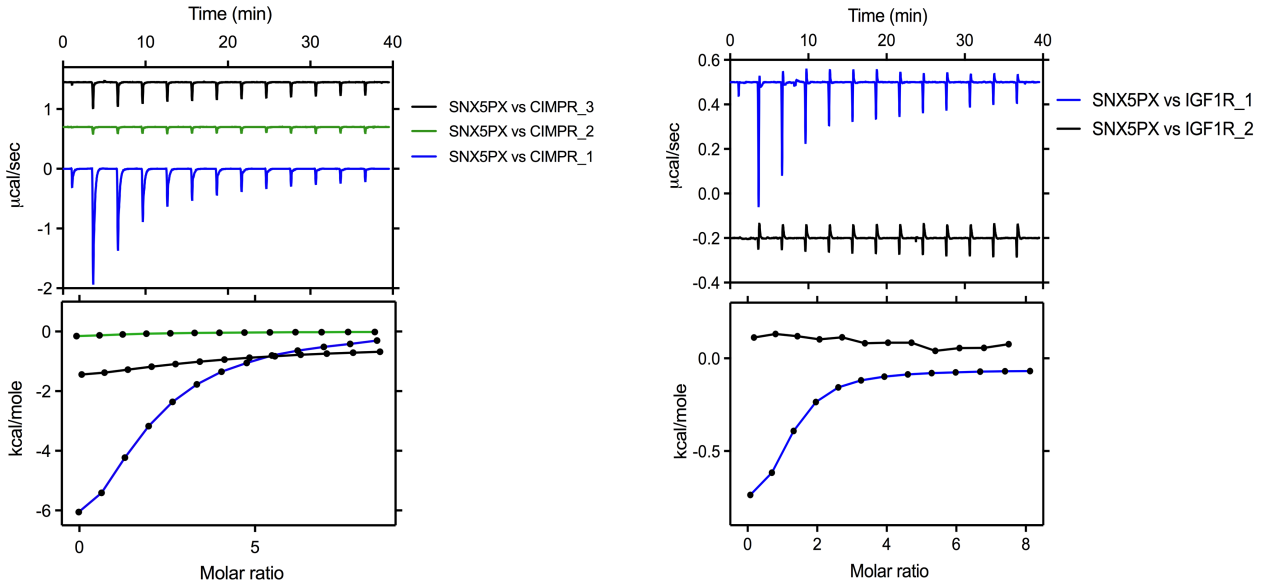


Figure 37. *ITC profiles for the binding of the PX domains of SNX5, SNX6 and SNX32 to the transmembrane cargoes (**A**) CIMPR_1 (SNVSYKYSKVNKEEETDENETEWLMEEIQ), (**B**) IGF1R_1 (GFREVSFYYSEENKL) and (**C**) SEMA4C (WDPVGYYYSDGSLKIVPGHARCQPG).*

To validate the region of binding on the transmembrane cargoes, I tested the interaction abilities of mutated residues in the hydrophobic patch on CI-MPR_1 and

IGF1R_1 as well as the WLM motif in CI-MPR_1. CI-MPR_1 mutations in the V-x-F-x-Y-S motif (CI-MPR_2) blocked binding to the SNX5PX, whereas mutations to the WLM motif (CI-MPR_3) reduced the affinity but still bound weakly to SNX5PX (**Fig. 38A**). The IGF1R_1 mutations in the V-x-F-x-Y-Y motif (IGF1R_2) also failed to bind to SNX5PX (**Fig. 38B**), suggesting that the hydrophobic motif in the cytoplasmic tail of the cargoes are important in the interaction with the SNX-BAR proteins. The SNX5PX mutant F136D was also tested against CIMPR_1 to confirm the importance of this residue in SNX5PX for the binding. CIMPR_1 was completely inhibited from binding to SNX5PXF136D (**Fig. 37A**).



A **B**

Figure 38. *ITC profile for the binding of the SNX5PX domain to the CIMPR_1 and IGF1R_1 mutants; (A) CIMPR_2 (SNVS**N**KYSKVNKEEETDENETEWLMEEIQ) and CIMPR_3 (SNVSYKYSKVNKEEETDENETE**AAA**EEIQ), and (B) IGF1R_2 (RE**A**S**A**Y**A**SEENKLPE).*

4.3.3 Purification of native and seleno-methionine SNX5PX-CIMPR fusion protein for crystallisation

In order to study the structural interaction between SNX5PX and the cytoplasmic tail of CIMPR through X-ray crystallography, I successfully expressed and purified SNX5PX-CIMPR fusion constructs as described under 4.2.2. Samples from every step of purification were analysed on SDS-PAGE to ensure the presence of the protein of interest. **Fig. 39** shows the purification of the SNX5PX-CIMPR construct after the removal of the GST tag using a Superdex 200 column.

The theoretical molecular weight of untagged SNX5PX-CIMPR fusion is around 20.6 kDa and was eluted as dimeric species (when compared to the column standards) between 50 mL and 60 mL on a Superdex 200 (16/600) column. The purification yielded recombinant proteins of 27 mg/ml.

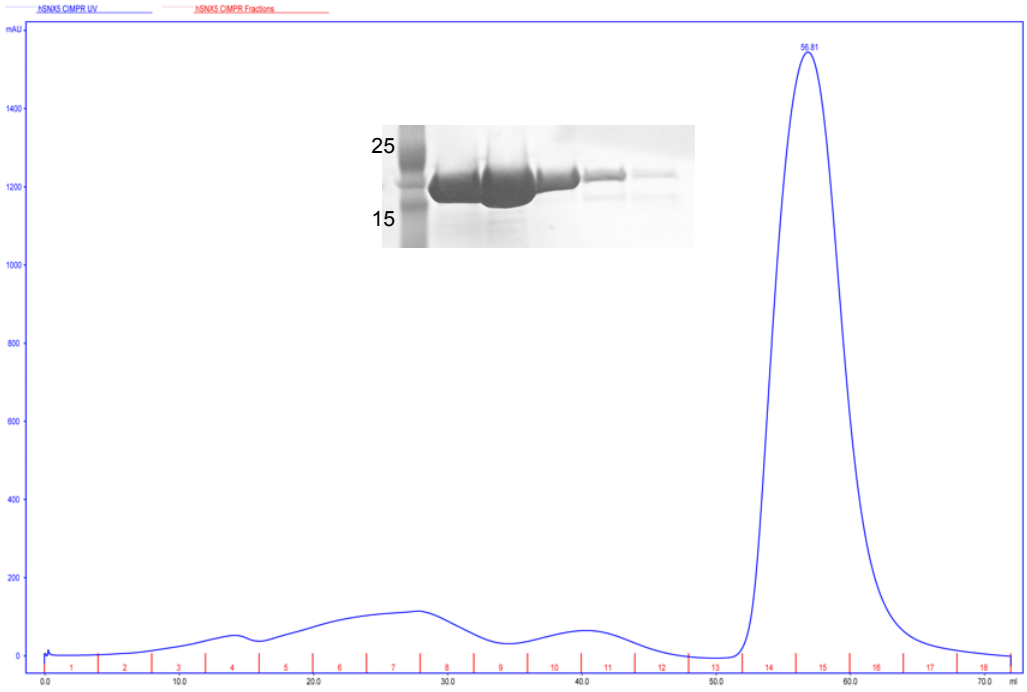


Figure 39. *FPLC chromatography profile of untagged SNX5PX-CIMPR fusion. The major peak (between 50-70 mls) fractions (13-17) correspond to the single bands on the gel (20.6 kDa).*

To allow me to unambiguously assign the binding site of the CIMPR sequence I also expressed and purified seleno-methionine SNX5PX-CIMPR fusion protein. As the CIMPR sequence includes a Met residue the anomalous signal from the incorporated selenium atom should allow me to identify its precise binding site even with low-resolution data. The seleno-methionine expression was kindly performed for me by Mr. Ryan Hall who is a research assistant in our lab, whereas the purification and crystallisation trials were carried out by me. The purification was performed using 50 mM HEPES pH 7.5, 100 mM NaCl, 2 mM DTT using a Superdex 200 (16/600) column. **Fig. 40** shows the purification of seleno-methionine SNX5PX-CIMPR fusion. The purification yielded soluble proteins of 16 mg/ml.

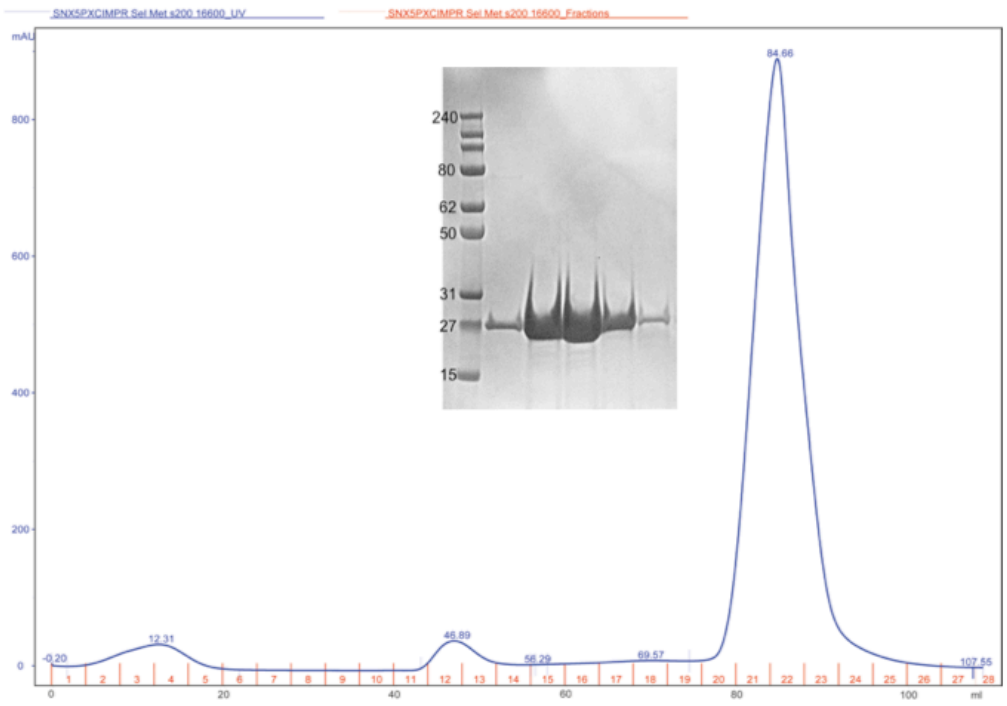


Figure 40. *FPLC chromatography profile of untagged seleno-methionine SNX5PX-CIMPR fusion. The fractions (20-24) from the major peak between 80-100 mls correspond to the monomeric species and the single bands on the gel.*

4.3.4 Crystallisation of native SNX5PX-CIMPR fusion complex and structure determination

96-well crystallisation screens of the native SNX5PX-CIMPR fusion yielded many spherulites. A few needle-shaped and rod-shaped crystals were formed under four conditions out of nearly 600 initial screen conditions that were set up using commercial kits. The needle crystals were sent to MX2 beamline to obtain the initial diffraction data (**Fig. 41A**). The reservoir condition that yielded the diffracting crystal was 30% (v/v) PEG 300, 0.1 M BIS-TRIS Propane pH 7.0, 15% (w/v) PEG 1000.

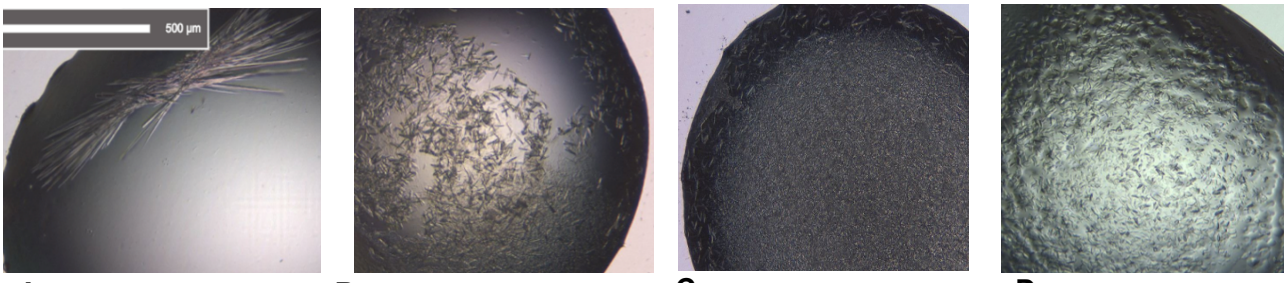


Figure 41. *Crystals of native SNX5PX-CIMPR fusion from 96-well plates. (**A**) Needle-crystals that were sent to MX2 beamline for the initial diffraction data. (**B-D**) Crystal hits of native SNX5PX-CIMPR in other reservoir conditions.*

After collecting low-resolution diffraction data, I have determined the structure of the native SNX5PX-CIMPR complex using molecular replacement, with the SNX5PX structure as the initial template model. The small crystals of the native SNX5PX-CIMPR have to date only diffracted to a low resolution of 3.1 Å and belong to spacegroup C2221. The C terminal of SNX5PX is connected to the N-terminal of CIMPR, and I could determine from difference electron density maps that the CIMPR sequence binds to the next SNX5PX in the asymmetric unit (**Fig. 42**). The binding site of CIMPR on SNX5PX is identical to where IncE binds to SNX5PX, and the density clearly reveals the formation of a similar β-hairpin structure. This provides structural evidence for the claim that IncE displaces CIMPR inorder to bind to SNX5PX.

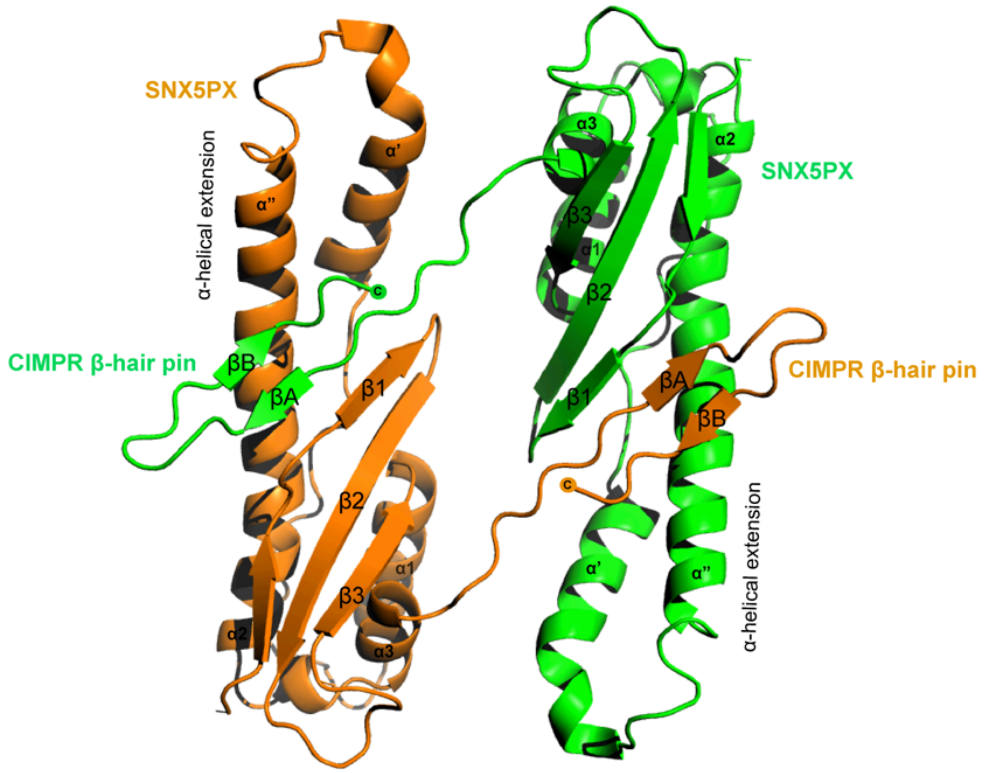


***Figure 42**. Native SNX5PX-CIMPR fusion crystal structure showing the CIMPR peptide (orange) bound to SNX5PX (green) as β-hairpin structure.*

The CIMPR forms a long β-hairpin structure that binds within a complementary hydrophobic groove, at the base of the unique extended helix-turn-helix insert between the ΨPxxPxK loop and α3 helix of SNX5PX, and adjacent to the β-sheet sub-domain (β1, β2 and β3) (**Fig. 42**). The main chain hydrogen bonds between CIMPR and SNX5PX are visible, but currently the resolution of the crystal structure is too low and the electron density is of relatively poor quality (**Fig. 43**), which prevents me from confidently assigning the CIMPR side chains. Thus from these initial small crystals I am currently unable to

precisely define the residues that are most important in forming a stable complex with SNX5PX. Further optimizations of the SNX5PX-CIMPR crystal hits are being undertaken, and I have also successfully reproduced the crystals using SeMet-labeled protein. These latter crystals will hopefully allow precise placement of the CIMPR "WLM" sequence within the complex crystal structure.

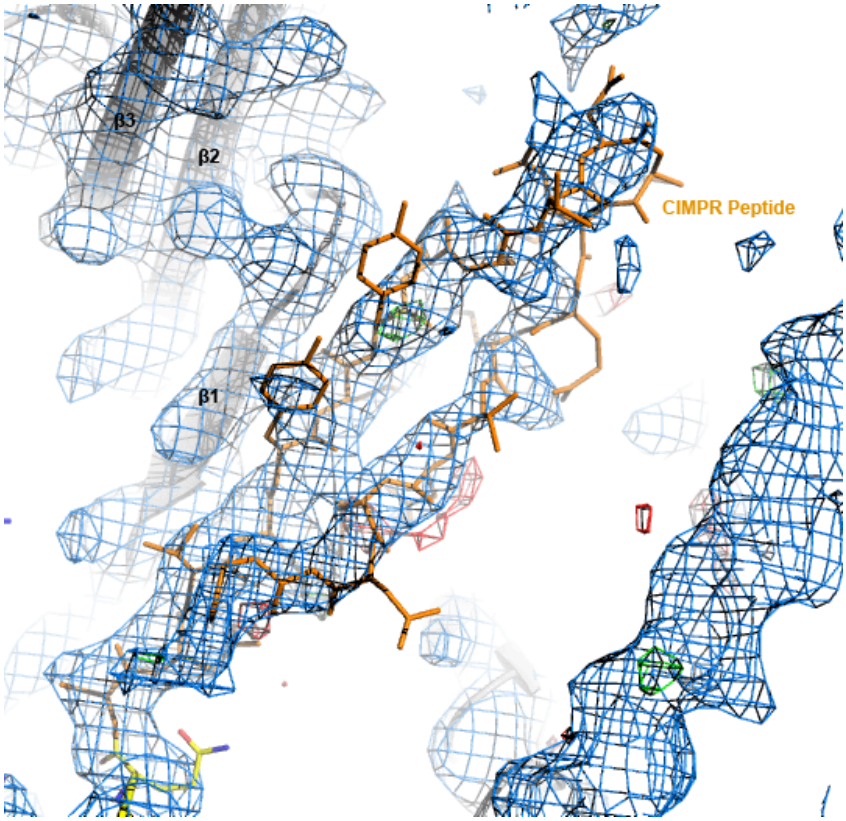
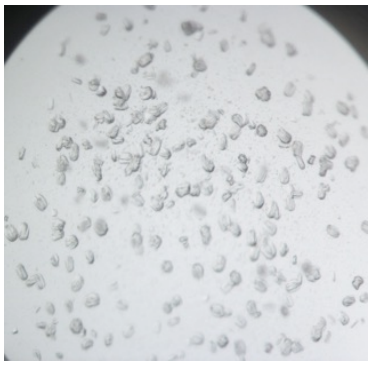


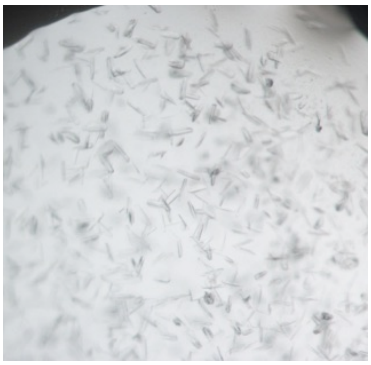
Figure 43. *CIMPR peptide (orange) modelled into the electron density.*

The needle crystals of the native SNX5PX-CIMPR from the 96-well plate that produced the above structure is now reproduced in a 24-well sitting drop crystallisation plate (**Fig. 44A**). Seleno-methionine SNX5PX-CIMPR crystals are optimised as well. Sel-Met SNX5PX-CIMPR produced initial hits in 96-well as clustered crystals in the reservoir condition 30% (w/v) PEG 5000 MME, 0.1 M MES pH 6.5, 0.2 M ammonium sulphate (**Fig. 44B**). This is now further optimised through addition of different concentrations of glycerol (**Fig. 44C**) and additive screen (**Fig. 44D, 44E** and **44F**), which has yielded promising crystals. These will be used to collect X-ray data.

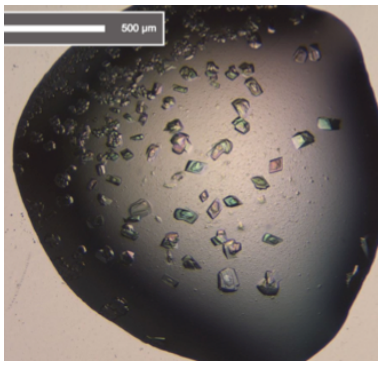
A



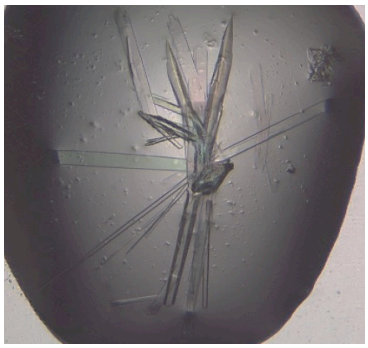
B



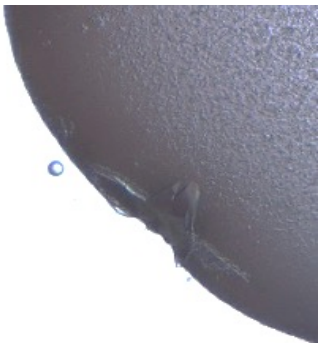
C



D



E



F

Figure 44. *(**A**) Optimised crystals of native SNX5PX-CIMPR in 30% (v/v) PEG 300, 0.1 M BIS-TRIS Propane pH 7.0, 15% (w/v) PEG 1000; (**B**) Clustered crystals of seleno-methionine SNX5PX-CIMPR in 30% (w/v) PEG 5000 MME, 0.1 M MES pH 6.5, 0.2 M ammonium sulphate (**C**) Clustered crystals optimised in 75% of 30% (w/v) PEG 5000 MME, 0.1 M MES pH 6.5, 0.2 M ammonium sulphate with 15% glycerol (**D**-**F**) Clustered crystal better with the addition of additive screen 40% (v/v) Polypropylene glycol P 400, 30% (v/v) (+/-)-2-Methyl-2,4-pentanediol and 50% (v/v) Jeffamine M-600 pH 7.0*

4.4 Discussion

Two recent studies (115,116) reported that transmembrane cargoes depend on the direct interaction with the heterodimeric combinations of SNX1 and SNX2 with the SNX5 and SNX6 proteins for their retrograde and endosomal recycling, independent of the Vps core of retromer complex. This is in contrast to the previously accepted models that proposed the Vps core subcomplex makes a direct contact with the transmembrane cargoes (119,120). Although the newest studies strongly suggest SNX5, SNX6 (and potentially

SNX32) are the cargo-interactive regions; it was unclear how this binding between SNX-BAR proteins and the cargoes are achieved. In this work I provide the first biochemical and structural evidence for the direct interactions of the SNX-BAR proteins with a specific cargo-sorting motif.

My previous SNX5PX-IncE structure and biochemical experiments (113), the comparative proteomic reports based on the mutations of the SNX5PX-IncE binding residues (115,116), and the displacement experiment reported by Elwell *et al*. (124), all lead to the hypothesis that CIMPR, IGF1R and SEMA4C binds to a common binding region within the PX domain of SNX5. In this chapter, using biochemical experiments and X-ray crystallography, I have demonstrated that the cytoplasmic tails of transmembrane cargoes CIMPR, IGF1R and SEMA4C, containing a conserved V-x-F/Y-x-Y-S region, directly interact with the PX domains of SNX5, and its homologues SNX6 and SNX32. Using ITC, the binding affinity of each cargo for the SNX-BAR proteins was measured. When compared to SNX6 and SNX32, SNX5PX bound tightly to the transmembrane cargos trafficked from endosome to the TGN; CIMPR_1, more than to the two other endosome to plasma membrane cargoes IGF1R_1 and SEMA4C. While SNX5PX and SNX32PX successfully interacted with all three cargoes, SNX6PX failed to bind to IGF1R_1. This has also been observed in unpublished immunoprecipitation data (Boris Simonetti and Peter Cullen, *personal communication*). This indicates a specific and likely subtle structural motif in SNX5 and SNX32 enables the binding of IGF1R_1, which is absent or hindered in SNX6PX. As expected, the ITC experiments using SNX5PXF136D mutation was found to block interactions with the cargoes. Mutations to the V-x-F/Y-x-Y-S motif and WLM motif of CIMPR as well as the V-x-F/Y-x-Y-S motif of IGF1R cargoes also confirms that there are critically important residues within the cargo motifs that are essential for the interaction.

The low-resolution crystal structure of SNX5PX-CIMPR fusion demonstrated and provided structural evidence that the CIMPR and IncE binds on the same binding site in SNX5PX. Even though CIMPR does not possess sequence similarity with IncE, the identified hydrophobic patch in SNX5 is enough for the binding to the CIMPR peptide. The poor electron density due to the low-resolution prevents me from confidently modelling the side chains at present and thus prevents me from recognising the specific residues in CIMPR important in forming a stable complex with SNX5PX. At present I believe the β-hairpin structure formed by CIMPR is likely to be composed of the N-terminal VSYKYS (strand 1) motif and the C-terminal WLM motif (strand 2). I hope to confirm this in future work by obtaining a higher resolution crystal structure, and I will also use seleno-

methionine labelling of the SNX5PX-CIMPR fusion to unambiguously place the WLM motif in the correct position within the structure. IGF1R and SEMA4C consist exclusively of the V-x-F/Y-x-Y-S region and do not possess the additional WLM motif. This suggests they could be potentially forming a single stranded linear structure through their V-x-F/Y-x-Y-S region rather than a β-hairpin structure. The crystal structures of SNX5PX with these different cargoes will enable us to understand the different conformations cargoes adopt to bind to SNX5PX. Structure-based mutagenesis will be performed to confirm the cargo interaction mechanisms *in vitro*, and can be tested for defective cargo trafficking *in vivo* in the future.

The continuous inflow of cargoes through endocytic routes and transport to different intracellular destinations contribute to the complexity of endosomal sorting pathways. Different reports of SNX family members as cargo-selective proteins have been published, and they have been linked to sorting of cargoes from endosomes to both the trans-Golgi network (TGN) and the plasma membrane. SNX3 proteins mediate the retrograde trafficking of metal transporter Dmt1-II and Wntless receptor (125,126). The PDZ domain of SNX27 forms a complex with the retromer subunit Vps26, and recognises specific recycling cargoes in phosphorylated states (127,128).

In our case, SNX-BAR proteins, SNX5, SNX6 and SNX32 can regulate endosome to TGN and endosome to plasma membrane trafficking of cargoes in a sequence-dependent manner. These diverse recycling pathways logically bring up a challenging question of how one protein complex can be involved in trafficking different cargoes to their different destinations. There are several possibilities that can be speculated on. One would be the formation of specific protein machineries around the SNX-BAR dimer sub complex that provide pathway specificity for the different cargo molecules. Kvainickas *et al.* (2017) propose a two-step procedure for the recycling of cargoes. The first step involves SNX-BAR complex binding specifically to the cargoes, and diverting them into a SNX-BAR tubule thus avoiding degradation of the cargoes. During the second step the additional protein machineries such as SNX3-retromer or SNX27-retromer then sort cargoes to other destinations

Understanding the different retromer-SNX cargo sorting pathways is not only important for understanding fundamental cellular physiology, but also has important implications in neurological diseases. Mutations in Vps35 are known to cause Alzheimer's and Parkinson's disease (20,129,130), which are found to cause accumulation of cargo in endosomes thus connecting neurological diseases to defects in endosomal sorting events. A complete understanding of endosomal sorting pathways and their functions will enable

us to work with selected trafficking pathways and are likely to open up new opportunities for the design of therapeutic interventions in the future.

Chapter 5

Chapter 5

Designing cyclic peptide inhibitors of the SNX-BAR proteins based on the *C. trachomatis* IncE structure

5.1 Introduction and significance

Cyclic peptides are chains of polypeptides that have cyclic ring structure, which are formed by linking through four possible ways; the N-terminal end of the peptide to the C-terminal end (head-to-tail), side chain-to-side chain, side chain-to-tail or head-to-side chain (**Fig. 45**). It has been shown that using cyclised peptides mimicking fragments of native structures can lead to the mimicking of their functions, and this has contributed to an increased interest in recreating different functional polypeptide regions to produce peptidomimetic molecules that are biologically active. Some of the initial targets for this interest were towards β- and γ-turns (131-136). But β-strands (137), β-sheets (138) and α-helices (136) are the most recent targets, which consist of polar NH and CO groups within the peptide backbones forming hydrogen bonds to promote and stabilize protein secondary structure.

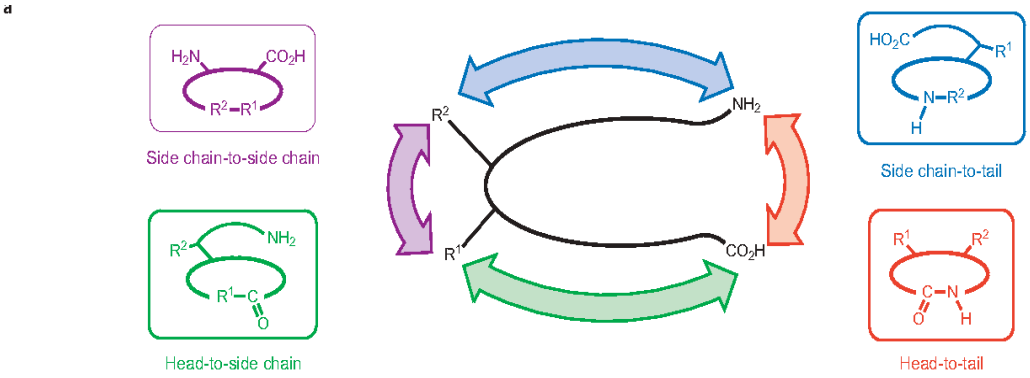


***Figure 45**. Four possible ways of constraining a peptide through macrocyclization (Figure adaped from (139).*

Even though peptides demonstrate various biological properties, they have been typically considered to be relatively poor drug molecules. The reasons include i) poor oral absorption ii) rapid metabolic turnover due to degradation by proteolytic enzymes and, iii) poor membrane permeability leading to failure of the peptide drug in reaching intracellular target. When compared to linear peptides however, cyclized peptides demonstrate

superior biological activity and proteolytic stability because of their conformational rigidity (140,141). The modification of biologically active peptides by introducing structural constraints is the most common way of producing cyclic peptides. Cyclic peptides have decreased entropy due to their rigidity, which often manifests as a significant increase in binding affinity with the target molecules. Cyclization also provides resistance to proteolytic degradation by removing both the free amino and carboxyl termini. Many cyclic peptides can also successfully cross the cell membrane, although just because a peptide is cyclised does not necessarily guarantee it will possess this property (142).

My studies described in chapter 2 and chapter 3 defined the molecular basis for how the chlamydial effector protein IncE binds and hijacks the SNX5 related proteins during chlamydial infection. The results demonstrated that the C-terminal binding sequence of IncE binds to the PX domain of the SNX5 and its homologues SNX6 and SNX32, forming an extended β-hairpin consisting of two antiparallel β-strands. Connected through a β-turn, the two IncE β-strands come together so that their N- and C-terminal residues are positioned very close together (**Fig. 46**). My work in Chapter 4 showed that the IncE peptide is mimicking a native receptor interaction with SNX5, albeit with significantly higher affinity and using a different sequence. I hypothesized that the two IncE β-strands could be linked to form a cyclic peptide that might display superior binding and stability to the linear IncE molecule, and thus provide a tool that could be used acutely block normal SNX5/6/32 cargo interactions.

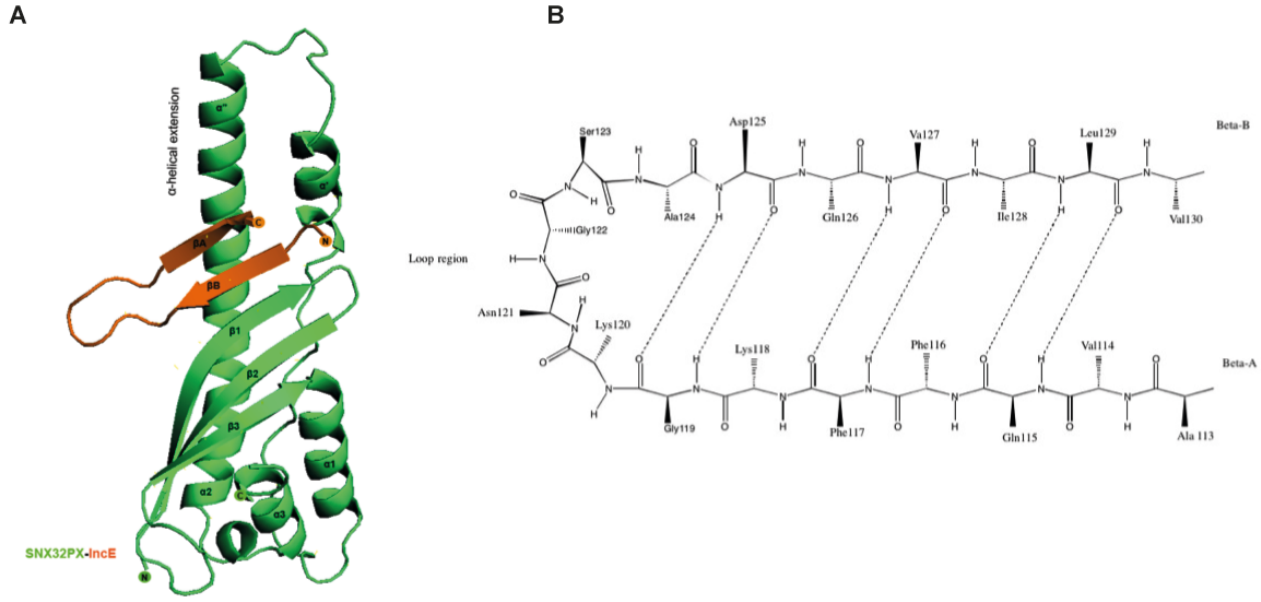


Figure 46. ***(A)*** *IncE (orange) bound to SNX32PX (green) domain as a β-hairpin;* ***(B)*** *IncE peptide with two antiparallel β-strands that is connected through a β-turn and an array of interstrand hydrogen bonds to stabilize the structure.*

In collaboration with Prof. David Fairlie and Dr. Timothy Hill (IMB, University of Queensland), around 20 linear and cyclic peptides derived from the sequence of IncE were engineered to explore the idea that a structurally constrained cyclic peptide could be designed to bind to SNX5, SNX6 and SNX32 with an increased affinity due to its higher rigidity and lower entropy. The peptides that were synthesized included truncations of IncE to shorten the extended loop, removing residues from the peptide that do not make a direct interaction with SNX5 based on our crystal structure, as well as different linkages between N- and C- termini to form cyclisation (**Table 6**).

In cyclic peptides, turns are motifs that reverse the direction of peptide strands and are classified as γ, β, or α depending on the increasing size of the turns. Even though cyclisation of peptides is the most commonly used method to stabilize turn structures, it is often not sufficient by itself and hence other turn-inducing structural motifs are introduced for further cyclic peptide constraints. In the initial round, some of the β-turn inducers were introduced to the IncE linear peptide to explore if those helped in forming a stable β-turn. Out of all amino acids, proline forms reverse turns in nature. Thus incorporating L-proline (peptides **1-3**) or heterochiral sequences such as D-proline-L-proline (peptides **4-10,17-18**) into linear peptides introduces turn-inducing constraints. Heterochirals have been previously shown to have the ability to be strong nucleators of β-hairpins (143). From **peptide 12-16** and **19-21**, besides β-turn inducers, cyclisations of the linear peptides were also introduced for the further enhancement of constraints of the peptide and to explore whether cyclisation increased the affinity of the peptides. **Peptides 11-14** contain a hairpin sequence of D-Proline-L-Proline for head-to-tail cyclisation to mimic a constrained cyclic peptide β-turn. This cyclisation approach has been introduced onto peptides with and without semi rigid β-turn inducers within the loop such as L-proline and D-proline-L-proline heterochirals. The other amino acid tail to sidechain cyclisation approach is a Lys-Asp lactam linker, which was introduced in **peptides 15-16** on top of D-Proline-L-Proline β-hairpin nucleator. The idea behind this in our case was to try and make a less constrained cyclic peptide to see if this had any effect on activity. **Peptides 19-21** contained δ-linked ornithine ($^{\delta}$Orn), which mimics a β-turn in a β-hairpin and connects the N-and C-termini.

Previously in chapter 2 and 3, I have demonstrated that the IncE binds the SNX5 PX domain with a K_d of ~1.9 μM. In chapter 4, I have discussed that recent reports have confirmed that the IncE is able to displace and block the binding of SNX5 to its selected transmembrane cargoes thus disrupting its trafficking (116,124). The cyclic peptides in this chapter will allow me to produce higher affinity IncE that can compete with the cargo binding to SNX5 *in vitro*, and in the future if we can confirm an ability to cross the cell

membrane to block the cargo binding *in vivo*, thus enabling me to make a specific inhibitory tool to study the role of SNX5 in membrane trafficking.

In this chapter, using Isothermal titration calorimetry (ITC), I demonstrate and confirm the hypothesis that one of the cyclised IncE peptide binds to the PX domain of SNX5 with an increased affinity, which is nearly 30x more than the standard IncE peptide, and ~300x higher affinity than endogenous cargo molecules. This cyclised peptide was also able to inhibit the binding of a transmembrane cargo, CIMPR to SNX5 *in vitro*. These results are very promising, and I am currently optimising crystals of this cyclic peptide complex to analyse its specific binding mechanism using X-ray crystallography. This information will be used to fine-tune this cyclic peptide to have even higher affinity and solubility, and the ability to cross the cell membrane thus leading to my final goal.

5.2 Materials and methods

5.2.1 Peptides

All linear and cyclic IncE-derived peptides used for ITC and crystallisation experiments were synthesised by Dr. Timothy Hill from Prof. David Fairlie's laboratory (IMB, UQ) (**Table 6; at the end of the section 5.2**). For the ITC experiments, peptides were weighed and dissolved in 50 mM HEPES (pH 8.0) and 100 mM NaCl (ITC buffer) to make a stock peptide concentration of 2 mM, which was diluted to 0.75 mM before use. The bacterial expression construct for human SNX5 PX domain proteins (SNX5 residues 26–170) were synthesized and cloned into pGEX-4T-2 by Genscript (USA). The constructs are N-terminally GST-tagged with site- specific thrombin cleavage sequence located between the GST moiety and the protein of interest. The CI-MPR peptide was purchased from Genscript (Hong Kong).

5.2.2 Isothermal titration calorimetry

ITC experiments were performed on a MicroCal PEAQ-ITC instrument at 25°C. The proteins were buffer exchanged into ITC buffer (50 mM HEPES (pH 8.0) and 100 mM NaCl) by gel filtration prior to ITC experiments. Linear and cyclic peptides at 0.75 mM were titrated into 0.05 mM SNX5PX in the cell. The binding data was processed using MicroCal PEAQ-ITC analysis software with a single site-binding model to determine the stoichiometry (n), the binding constant (K_d), the estimated heat of binding (ΔH), the Gibbs free energy (ΔG) and the binding entropy (ΔS). Three experiments were performed for each set of samples to determine the average ± standard error of the mean (SEM) for thermodynamic quantities, except for the linear peptides and the ones that did not interact

with SNX5PX experiments as well as the control, where only single experiments were performed. For these single experiments, all experiments were performed using the same batch of protein to allow direct comparisons to be made.

5.2.3 Crystallisation of cyclic peptide complexes

SNX5, SNX6 and SNX32 PX domains were concentrated to 12 mg/ml for crystallization with cyclic peptide **12** (**Table 6**). These proteins were directly mixed together with the appropriate peptide at a 1:2 molar ratio of protein to peptide and were incubated on ice for 1 h. Three 96-well crystallization hanging-drop screens for each experiment were set up using a Mosquito Liquid Handling robot (TTP LabTech) at 20°C in the UQ ROCX facility. After setting up these 96-well screens, the trays were incubated at 20°C in a Rockimager storage hotel (Formulatrix), where the drops were imaged at different time points for 21 days.

Table 6. List of linear and cyclic peptides

C	PAVQFFKGKNGSADQVILV	Control
1	PAVQFFKGKN**P**SADQVILV	Linear
2	PAVQFFKGK**P**GSADQVILV	Linear
3	PAVQFFK**P**KNGSADQVILV	Linear
4	PAVQFFKGK**pP**SADQVILV	Linear
5	PAVQFFK**pP**NGSADQVILV	Linear
6	PAVQFFKGK**Pp**SADQVILV	Linear
7	SPANGPAVQFFKGK**pP**SADQVILVTQ	Linear
8	SPANGPAVQFFKG**pP**GSADQVILVTQ	Linear
9	SPANGPAVQFFK**pP**NGSADQVILVTQ	Linear
10	SPANGPAVQFFKG**Pp**GSADQVILVTQ	Linear
11	**P**AVQFFKGKDQVILV**p**	Head-to-tail cyclisation
12	**P**AVQFFKGKN**P**SADQVILV**p**	Head-to-tail cyclisation
13	**P**AVQFFKGKDQVILV**Pp**	Head-to-tail cyclisation
14	**P**AVQFFKGK**pP**SADQVILV**p**	Head-to-tail cyclisation
15	Ac**K**PAVQFFKGK**pP**SADQVILV**D**	Side chain-to-side chain cyclisation
16	Ac**K**AVQFFKGK**pP**SADQVIL**D**	Side chain-to-side chain cyclisation
17	KADQVILV**pP**AVQFFK	Linear
18	KDQVILV**pP**AVQFFK	Linear
19	AVQFFKGKDQVILV**O**	Head-to-side chain cyclisation
20	AVQFFKGKN**P**SADQVILV**O**	Head-to-side chain cyclisation
21	VAVQFFKGKN**P**SADQVILV**O**	Head-to-side chain cyclisation

P= L-Proline; **p**= D-Proline; **pP**=D-proline-L-proline heterochirals; **O**= δ-linked ornithine ($^{\delta}$Orn)

5.3 Results

5.3.1 Interaction studies of linear and cyclic peptides with SNX5PX using ITC

The numbering scheme for the IncE peptide is taken from the original IncE sequence from *C. trachomatis* serovar L2 (Pro112, Ala113, Val114, Gln115, Phe116, Phe117, Lys118, Gly119, Lys120, Asn121, Gly122, Ser123, Ala124, Asp125, Gln126, Val127, Ile128, Leu129, Val130) and is used for all the peptides in this chapter.

Control IncE peptide was used as the binding reference. This peptide demonstrated a K_d of 1.9 μM to SNX5PX (**Fig. 47A**). Peptides **1-10** and **16-17** were generated during the first round of linear IncE peptides that were synthesised. A single batch of SNX5 PX domain was tested for binding against all the linear IncE peptides.

Peptides **1-3** consist of all residues from Pro112 to Val130, with one residue in loop region (Gly119-Ala124) being replaced by L-proline at a time. Peptides **1-2** showed similar K_d (1.6 μM and 1.7 μM) to the control (**Fig. 47B** and **47C**), which showed replacement of either Gly122 or Asn121 with L-proline does not improve affinity, whereas results from peptide **3** (replacement of Gly119 with L-proline) are inconclusive due to the difficulty in dissolving the peptides in the buffer (**Fig. 47D**).

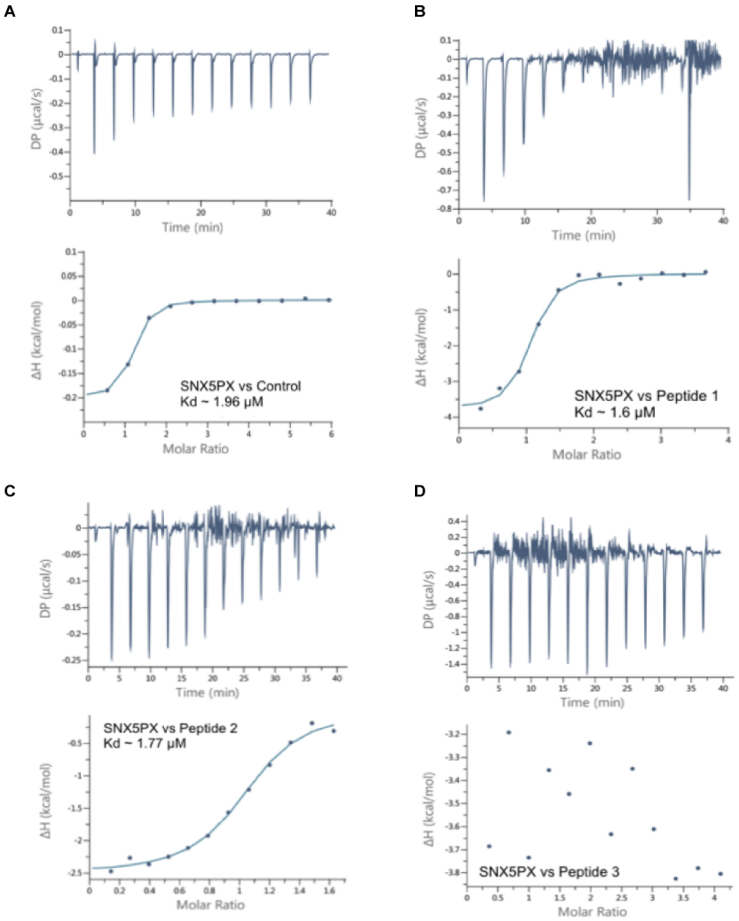


Figure 47. *ITC profiles for the binding of IncE linear peptides to SNX5PX domain.* ***(A)*** *Control (PAVQFFKGKNGSADQVILV),* ***(B)*** *Peptide* ***1*** *(PAVQFFKGKN**P**SADQVILV),* ***(C)*** *Peptide* ***2*** *(PAVQFFKGK**P**GSADQVILV) and* ***(D)*** *Peptide* ***3*** *(PAVQFFK**P**KNGSADQVILV) vs SNX5PX.*

Instead of L-proline, residues in the loop region in peptides **4-6** are replaced by D-proline-L-proline or L-proline-D-proline heterochirals. The loop region residues Asn121 and Gly122 are replaced with D-proline-L-proline in peptide **4**, which improved the K_d (1.07 μM) when compared to the binding reference (**Fig. 48A**). Peptide **5** has the D-proline-L-proline two residues downstream compared to peptide **4** at Gly119 and Lys120, which demonstrated no binding (**Fig. 48B**). Peptide **6** is similar to peptide **4** in regards to the residues present, except the D-proline-L-proline heterochirals at Asn121 and Gly122 are replaced with L-proline-D-proline (**Fig. 48C**). This reversal of prolines in peptide **6** did not favour an improved binding affinity (2.6 μM) more than peptide **4**.

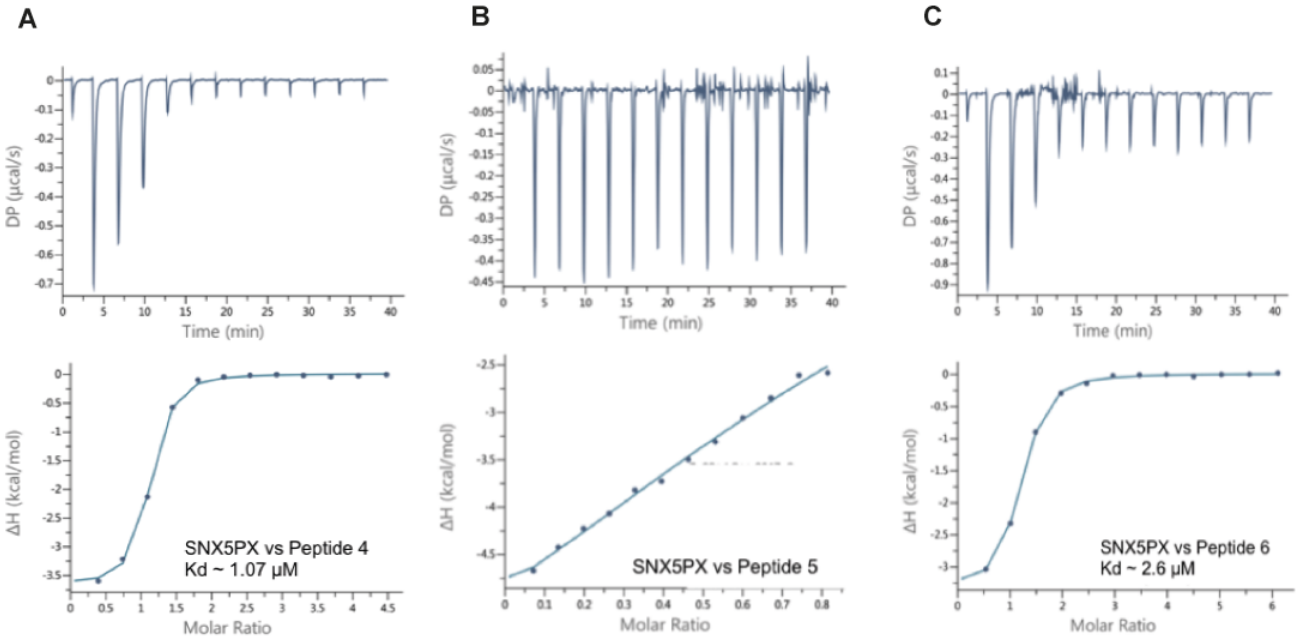


Figure 48. *ITC profiles for the binding of IncE linear peptides to SNX5PX domain.* ***(A)*** *Peptide* ***4*** *(PAVQFFKGK**pP**SADQVILV),* ***(B)*** *Peptide* ***5*** *(PAVQFFK**pP**NGSADQVILV), and* ***(C)*** *Peptide* ***6*** *(PAVQFFKGK**Pp**SADQVILV) vs SNX5PX.*

Peptides **7-10** have either D-proline-L-proline or L-proline-D-proline heterochirals at the loop region with extended IncE residues on the N- and C- termini (Ser107, Pro108, Ala109, Asn110, Gly111 from N-terminus and Thr131, Gln132 from C-terminus). Peptide **7** consists of D-proline-L-proline at Asn121 and Gly122 as in peptide **4**, but on top of that it possesses the extended N- and C- termini residues as well, which yielded identical K_d (1.09 μM) to peptide **4** **(Fig 49A)**. This demonstrates that these extended residues (Ser107, Pro108, Ala109, Asn110, Gly111 from N-terminus and Thr131, Gln132 from C-terminus) do not contribute towards any important interactions with SNX5 PX as demonstrated in Paul *et al*. (113), but the D-proline-L-Proline in the loop region is better

than L-proline alone (peptides **1-3**). Peptides **8** and **9** consist of D-proline-L-proline heterochirals one residue downstream of each other at positions Lys120, Asn121 and Gly119, Lys 120 respectively. Peptide **10** has the heterochirals reversed at Lys120 and Asn121. Peptides **8** and **10** demonstrated weaker binding affinity (31 μM and 7.7 μM respectively) (**Fig. 49B** and **49D**) than peptide **7,** whereas peptide **9** showed lack of binding. The binding affinity of peptide **9** is inconclusive because this peptide was not completely dissolvable in buffer (**Fig. 49C**). This result indicates even a simple shift in the replacement of residues can affect affinity and solubility of peptides. Thus out of all the linear peptides, peptide **4** and **7** (1.07 μM and 1.09 μM) were the best when compared to the K_d of reference peptide (**Fig. 48A** and **49A**).

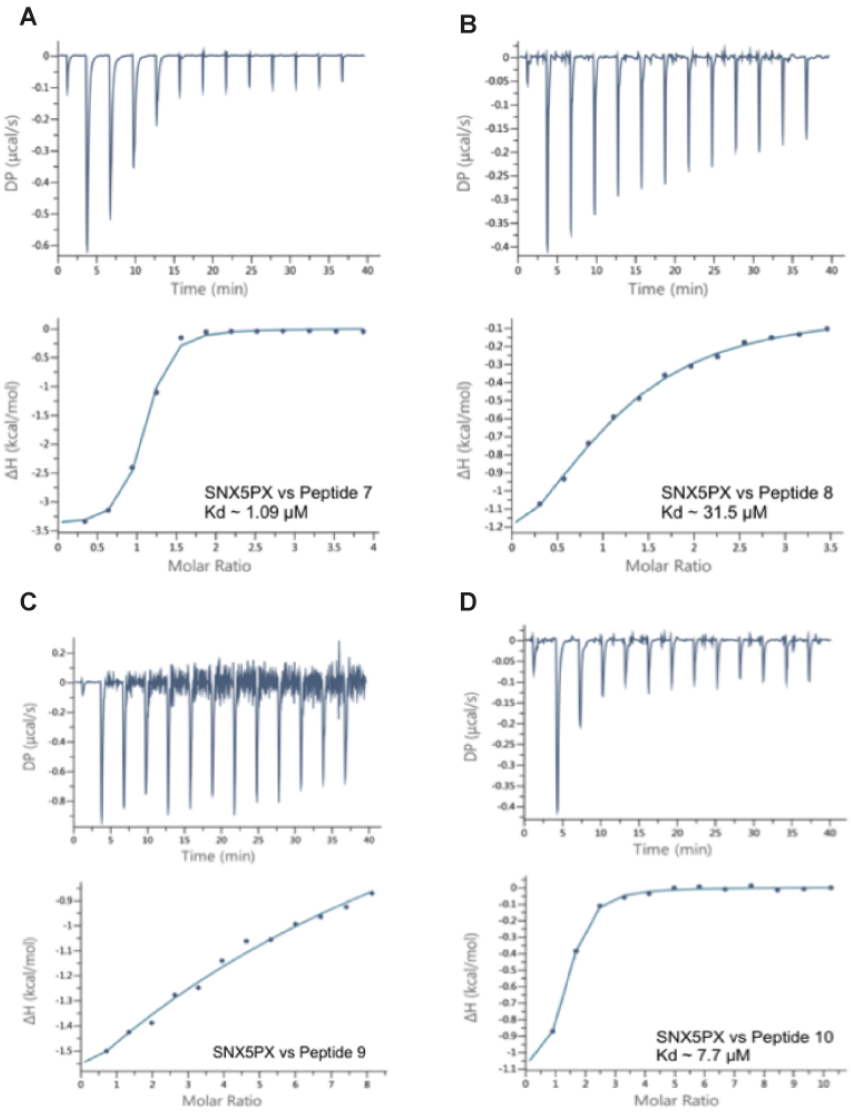


Figure 49. *ITC profiles for the binding of IncE linear peptides to SNX5PX domain.* ***(A)*** *Peptide* ***7*** *(SPANGPAVQFFKGK**pP**SADQVILVTQ),* ***(B)*** *Peptide* ***8*** *(SPANGPAVQFFKG**pP**GSADQVILVTQ),* ***(C)*** *Peptide* ***9*** ***(****SPANGPAVQFFK**pP**NGSADQVILVTQ), and* ***(D)*** *Peptide* ***10*** ***(****SPANGPAVQFFKG**Pp**GSADQVILVTQ) vs SNX5PX.*

Peptides **11-21** were cyclised to test the theory that the binding affinity could be increased compared to linear peptides. Cyclisation of linear peptides were implemented through L-proline-D-proline N-to C (head-to-tail) linkage (peptides **11-14**), Asp-Lys lactam side chain-to-side chain linkage (peptides **15-16**) and δ-linked ornithine ($^{\delta}$Orn) head-to-side chain linkages (peptides **19-21**). Peptide **12** (L-pro-D-pro head-to-tail linkage, with L-proline at position Gly122) demonstrated the highest affinity of 67 nM (**Fig. 50B**), whereas peptide **11** with L-proline-D-proline linkage, with no L-proline at position Gly122 showed much weaker binding (25 μM) (**Fig. 50A**). Peptide **14** was synthesised to determine if the affinity of peptide **12** can be improved with D-proline-L-proline heterochiral at the same loop region (Asn121, Gly122) instead of L-proline alone. This peptide yielded an improved affinity (0.6 μM) (**Fig. 50D**) over the linear reference but lower affinity than peptide **12**.

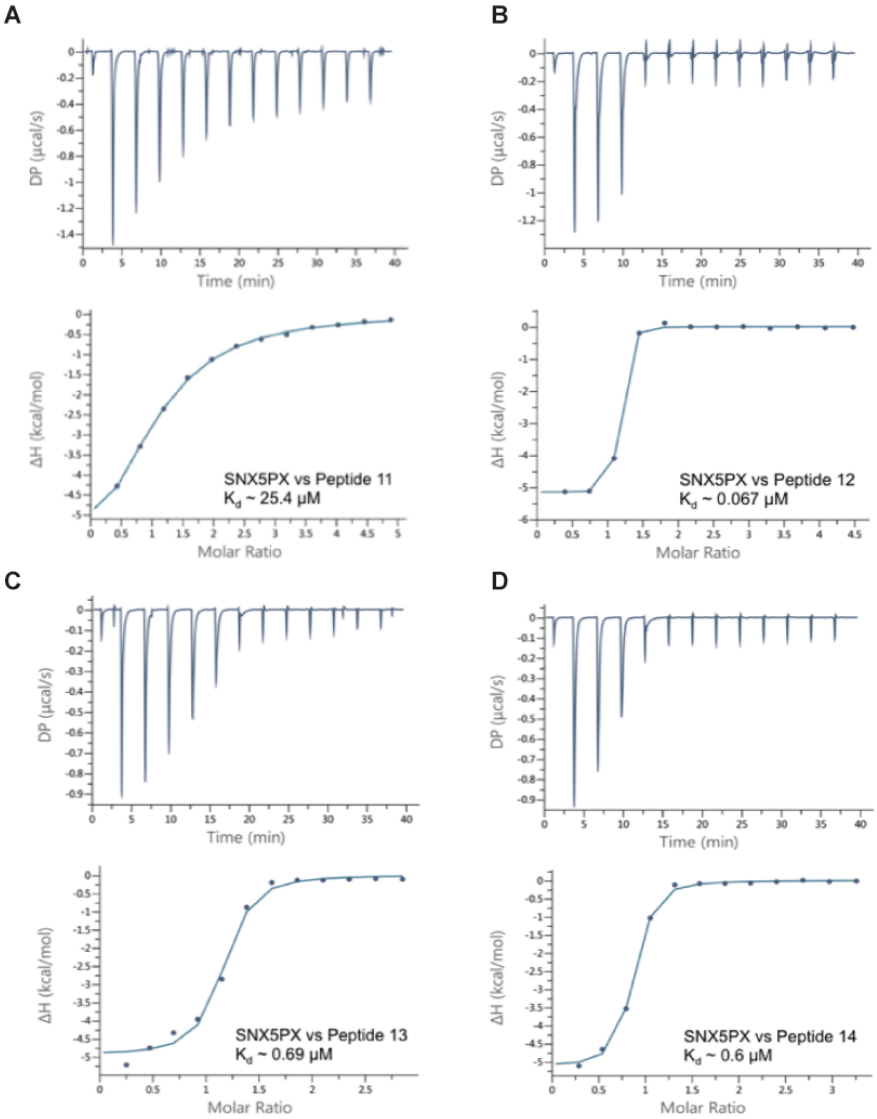


Figure 50. *ITC profiles for the binding of cyclic IncE peptides to SNX5PX domain.* ***(A)*** *Peptide* ***11*** *(**P**AVQFFKGKDQVILV**p**),* ***(B)*** *Peptide* ***12*** *(**P**AVQFFKGKN**P**SADQVILV**p**),* ***(C)*** *Peptide* ***13*** *(**P**AVQFFKGKDQVILV**Pp**), and* ***(D)*** *Peptide* ***14*** *(**P**AVQFFKGK**pP**SADQVILV**p**) vs SNX5PX.*

Cyclic peptides **15** and **16** with Asp-Lys cyclisation demonstrated a slightly better K_d than the reference peptide (1.8 μM and 1.54 μM respectively) (**Fig. 51A** and **51B**).

??

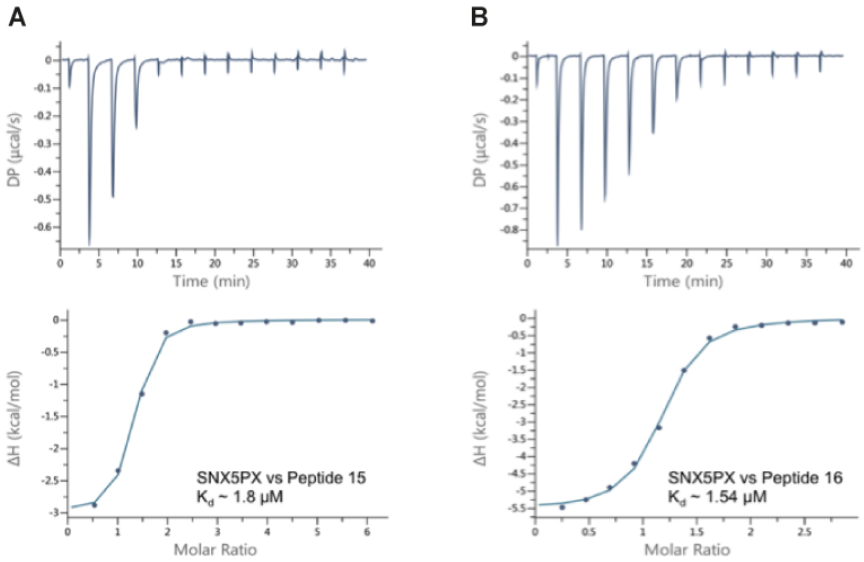


Figure 51. *ITC profiles for the binding of cyclic IncE peptides to SNX5PX domain.* ***(A)*** *Peptide* ***15*** *(Ac**K**PAVQFFKGK**pP**SADQVILV**D**) and* ***(B)*** *Peptide* ***16*** *(Ac**K**AVQFFKGK**pP**SADQVIL**D**) vs SNX5PX.*

Peptides **20** and **21** with ornithine linkages demonstrated weaker affinity of 16.6 μM and 7 μM (**Fig. 52B** and **52C**), whereas peptide **19** showed no binding (**Fig. 52A**). Since peptide **12** demonstrated nearly 30x more stronger affinity than the reference, this peptide was used for further *in vitro* inhibition experiments and crystallisation trials.

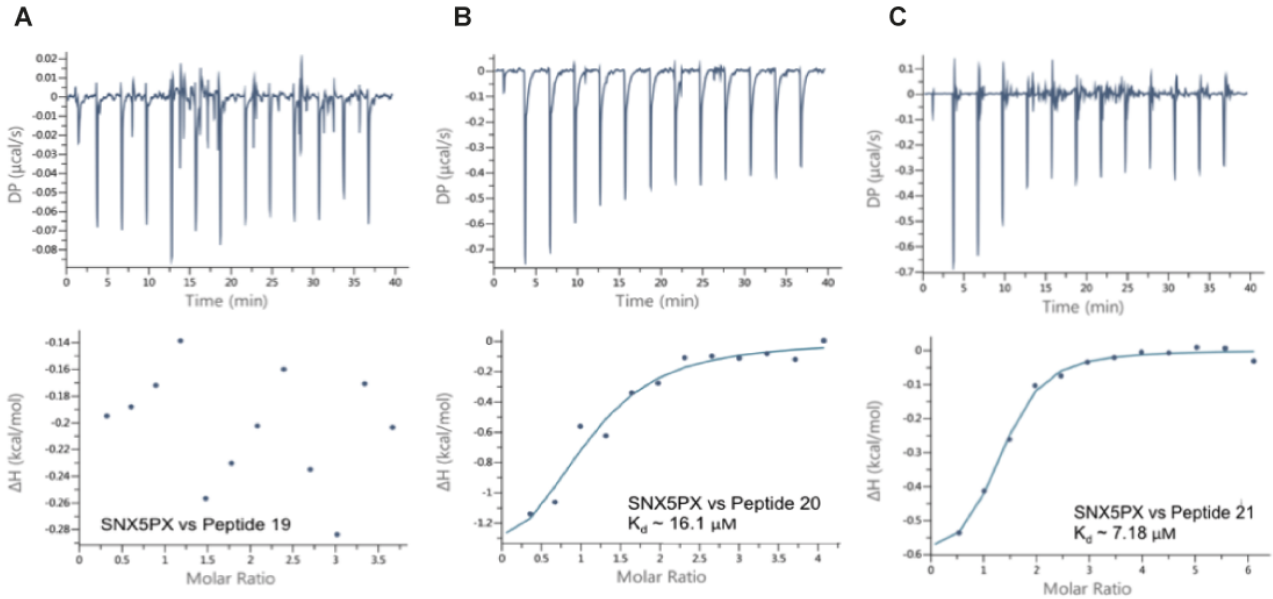


Figure 52. *ITC profiles for the binding of cyclic IncE peptides to SNX5PX domain.* ***(A)*** *Peptide* ***19*** *(AVQFFKGKDQVILV**O**),* ***(B)*** *Peptide* ***20*** ***(****AVQFFKGKN**P**SADQVILV**O**) and (C) Peptide* ***21*** *(VAVQFFKGKN**P**SADQVILV**O**) vs SNX5PX.*

Table 7. Thermodynamic parameters of the SNX5PX vs linear and cyclic peptides

Sample cell	Titrant	ΔH (kcal/mol)	TΔS (kcal/mol)	ΔG (kcal/mol)	K_d (μM)	N
SNX5 PX	Control	-0.202	-7.59	-7.79	1.96	0.999
	1	-3.81	-4.11	-7.91	1.59	0.953
	2	-2.52	-5.33	-7.85	1.77	1.02
	3	/	/	/	/	/
	4	-3.69	-4.46	-8.15	1.07	0.980
	5	/	/	/	/	/
	6	-3.37	-4.23	-7.60	2.68	1.03
	7	-3.44	-4.70	-8.14	1.09	0.965
	8	-1.92	-4.22	-6.14	31.5	1.02
	9	/	/	/	/	/
	10	-1.22	-5.75	-6.97	7.7	1.08
	11	-7.52	-1.25	-6.27	25.4	0.958
	12	-6.1 ± 0.9	-3.94 ± 0.68	-8.9 ± 0.94	0.067 ± 0.01	1.07 ± 0.05
	13	-5.26 ± 0.29	-3.16 ± 0.19	-8.42 ± 0.101	0.69 ± 0.09	1.09 ± 0.07
	14	-8.23 ± 0.28	-3.88 ± 1.21	-8.48 ± 0.05	0.6 ± 0.05	0.97 ± 0.15
	15	-3.52 ± 0.77	-2.45 ± 0.15	-7.81 ± 0.12	1.8 ± 0.21	1.06 ± 0.12
	16	-5.59 ± 0.7	-2.71 ± 0.78	-8.06 ± 0.33	1.54 ± 0.37	1.07 ± 0.04
	17	-0.704	-7.90	-7.20	5.28	1.02
	18	/	/	/	/	/
	19	/	/	/	/	/
	20	-1.78	-4.85	-6.54	16.1	1.01
	21	-0.644	-6.32	-7.02	7.18	1.19

ΔH – enthalpy, T- absolute temperature, ΔS –change in entropy, ΔG – Gibbs free energy, K_d – dissociate constant, N – stoichiometry

5.3.2 Binding inhibition of transmembrane cargo to SN5PX by cyclic peptide

Using ITC, a simple inhibition experiment using cyclic peptide **12** was performed. The reaction cell was filled with a mix of SNX5 PX with a linear peptide from the transmembrane cargo CIMPR, where cyclic peptide **12** was titrated into the cell. This cyclic peptide successfully displaces CIMPR and binds to SNX5PX with an apparent K_d of 125 nM, in agreement with the high affinity of the cyclic peptide **12** alone. Interestingly, the binding profile indicates an endothermic interaction occurs rather than the exothermic interaction of peptide **12** alone. This may be due to the pre-binding of the CI-MPR peptide

leading to the stabilisation of the bent SNX5 conformation being primed for cyclic peptide displacement, and a subsequent change in the apparent thermodynamics of peptide **12** binding. In reciprocal experiments, the CIMPR failed to displace the cyclic peptide, showing no significant binding to SNX5 when pre-incubated with the high-affinity peptide **12** prior to titration (**Fig. 53**).

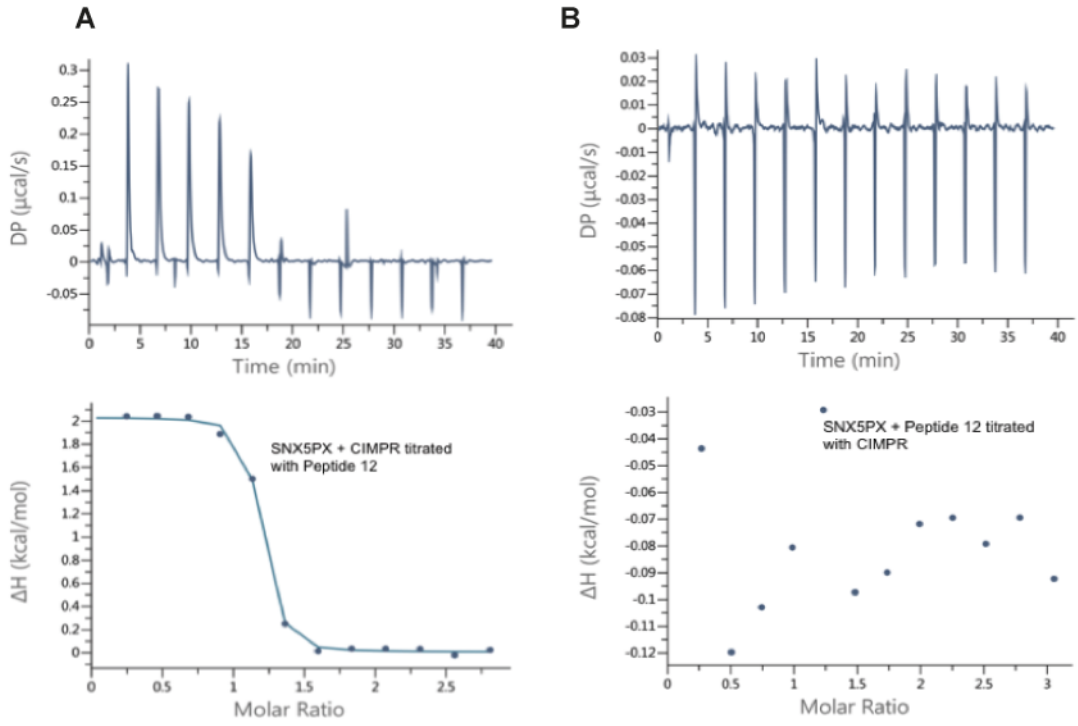


Figure 53. *ITC profiles showing the binding inhibition of transmembrane cargo, CIMPR by the cyclic IncE Peptide* ***12*** *(**P**AVQFFKGKN**P**SADQVILV**p**).* ***(A)*** *Peptide* ***12*** *titrated into cell containing SNX5PX and CIMPR and* ***(B)*** *CIMPR titrated into cell containing SNX5PX and Peptide* ***12****.*

5.3.3 Crystallisation trials of SNX5PX, SNX6PX and SNX32PX in complex with cyclic peptide

Out of a variety of initial crystallisation screens that I attempted in 96-well trays, one condition for SNX6 PX-cyclic peptide **12** and two for SNX32 PX-cyclic peptide **12** produced initial hits (**Fig. 54**). The reservoir condition that gave the initial hit of SNX6 PX-cyclic peptide **12** is 0.16 M calcium acetate, 0.08 M sodium cacodylate pH 6.5, 14.4% (w/v) PEG 8000 and 20% (v/v) glycerol. The first SNX32 PX-cyclic peptide **12** initial hit was obtained in 40% (v/v) PEG 600, 0.1 M sodium cacodylate pH 6.5, 0.2 M calcium acetate and the second hit was in the reservoir condition 0.2 M calcium chloride, 0.1 M

HEPES sodium salt pH 7.5, 48% (w/v) PEG 400. These conditions are currently being optimised to obtain macrocrystals large enough for a high-resolution structural analysis.

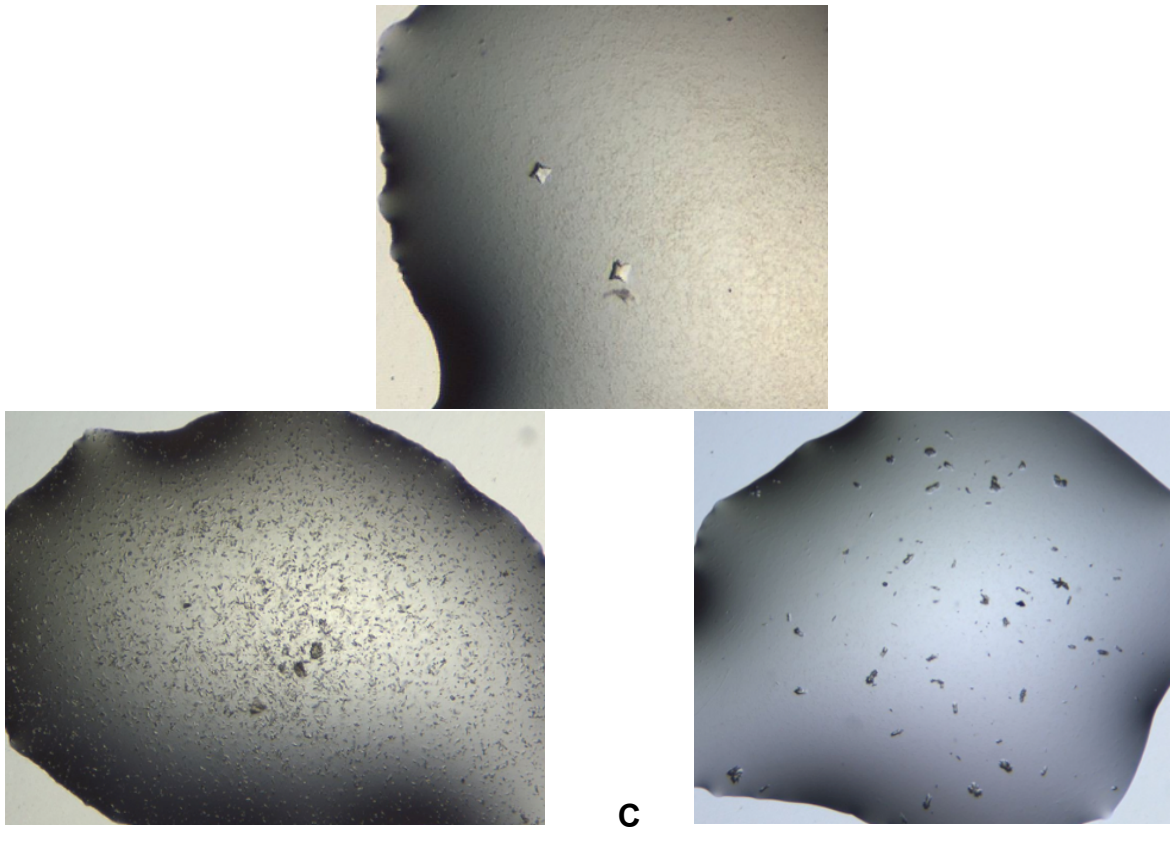


Figure 54. *Crystals of (**A**) SNX6PX-cyclic peptide **12** complex; (**B**-**C**) SNX32PX-cyclic peptide **12** complex.*

5.4 Discussion

In virtually all cellular processes, biological functions are influenced by protein–protein interactions (PPIs). These PPIs can be formed between two protein domains that are structured, a structured domain and a short peptide, or even between two peptides. One of the currently promising strategies in therapeutics is to modulate these PPIs (144), a strategy that also leads to the generation of reagents that can be used to better understand the biological processes being targeted (144-146). Peptides and peptidomimetics are an important class of molecules with particular relevance to modulating PPI interactions, especially when structural information can be used to guide the design of these PPI inhibitors. When separated from their protein-bound state, peptides generally will lose their secondary structure and conform to a flexible nature in solution. These flexible peptides are exposed to degradation through proteolysis, and also typically show lower affinity for their target. Thus mimicking the peptides in their most

efficient and bioactive conformation has been one of the goals for the successful designing of cyclic peptides as protein-protein interaction inhibitors.

The chlamydial effector protein IncE (residues 112-130) was found to form a β-hairpin structure when bound to SNX5 and SNX32. The N-terminal βA strand of the IncE sequence consists of Gly112-Lys118 and the antiparallel C-terminal βB strand of IncE is formed by Asp125-Val130. These two strands are connected by a short loop (Gly119-Ala124). The IncE βA strand forms main chain hydrogen bonds with the SNX5 PX β-sheet, whereas βB forms an interface with the α-helical region of SNX5. The IncE loop region does not make any direct contact with the SNX5 protein. In an attempt to explore the idea of deriving cyclic peptides from linear IncE sequences several different approaches were tried. As explained previously under section 5.1, these different approaches included introduction of different β-turn inducers, different cyclisation approaches, and different truncations of the βA, βB and the loop regions that potentially will affect the affinity and solubility.

Isothermal titration calorimetry was used to characterize the binding affinity of all the above-mentioned (**Table 6**) linear and cyclic peptides with the SNX5PX domain. Besides β-turn inducers and cyclisation approaches, the importance of IncE residues from βA and βB strands and loop regions were also investigated using truncations and replacement of residues to explore whether any of these approaches increase the affinity and solubility. Single residue positions in the loop region were screened by replacement with L-proline. Single replacements of Gly119, Asn121 or Gly122 in the loop region with L-proline did not make an improvement in K_d when compared to the control (K_d 1.9 uM). But replacements of both Asn121 and Gly122 with D-proline-L-proline heterochirals made approximately 2x improvement in K_d. On the other hand, switching of heterochirals to L-proline-D-proline at the same positions under the same conditions yielded weaker K_d. Introduction of D-proline-L-proline heterochirals at Gly119 and Lys120, as well as at Lys120 and Asn121 also suffered loss of affinity. Thus linear peptides with truncations of a few residues from the N and C termini and replacements of loop region residues with the D-proline-L-proline heterochiral β-turn inducer demonstrated to have improved affinity for SNX5PX. Cyclisation of linear peptides besides different β-turn inducers were found to improve the affinity significantly. The D-proline-L-proline and Asp-Lys lactam cyclisations together with the D-proline-L-proline β-turn inducers in the IncE loop region demonstrated similar K_d indicating similar contacts in the protein-peptide interface. Cyclised peptide (L-proline-D-proline) with deletion of a part of the loop region (Asn121, Gly122, Ser123, Ala124) as well as the introduction of L-proline on βB strand at Arg126 position also demonstrated similar

but slight weaker affinity. The same conditions without the L-proline or any other β-turn inducers demonstrated much weaker K_d of 25 uM, whereas re-introduction of the deleted loop region and L-proline at the Gly122 position yielded a K_d of 56 nM. The same condition, but the $^{\delta}$Orn linker only demonstrated a K_d of 16 uM. Thus a L-proline-D-proline cyclisation with the replacement of Gly122 loop residue to L-proline improved the binding of this IncE cyclic peptide with SNX5PX by 30x more than the control.

In the future, the crystal structure of the cyclic peptide in complex with SNX5PX or its homologues will enable us to understand and validate biophysical data using X-ray structural data. The structural data will confirm the backbone conformation and sidechain interactions established by the cyclic peptide to the protein, providing insights into the mechanism leading to higher affinity when compared to the linear peptide. The interactions formed between the cyclic peptide with the protein are expected to be similar to the interactions established by the linear IncE peptide with the protein however, and the biggest contributor to the increased affinity is likely to be pre-stabilisation of the peptide backbone due to residue substitutions that allows a better fit of the cyclic peptide in the protein binding pocket. It is possible that further modifications such as cross-linking of the β-strands might lead to even further improvements in affinity.

The ability of the higher affinity cyclic peptide **12** to inhibit binding of transmembrane cargo CIMPR to SNX5PX *in vitro* demonstrates the potential to take up this positive cyclic peptide hit as a starting point, to improve its affinity and solubility to use as an inhibitory tool. The next steps will be to determine if the peptide is able to cross the cell membrane. On top of cyclisation, certain structural features of a peptide are typically associated with an improved ability to cross the cell membrane. Cyclosporin A, a well-known immunosuppressant, is a good example of a cyclic peptide that is cell permeable. This cyclic peptide keeps the hydrophilic groups away from the surface of the molecule through several intra-molecular hydrogen bonds. Our future goals include testing the cell permeability of the cyclic peptides to explore whether they can cross the cell membrane and block the binding of transmembrane cargoes *in situ*, and if required working on the modifications of this cyclic peptide to increase its cell permeability without negatively affecting its solubility and affinity.

As indicated in the previous chapters, SNX-BAR proteins are emerging as key players in retromer-independent trafficking pathways. Understanding these pathways are important since the transmembrane cargoes that are trafficked by SNX-BAR proteins, are involved in cancer and neurodegenerative diseases. The new cyclic peptides can be developed to fine-tune the SNX-BAR mediated trafficking pathways by inhibiting its

interaction with cargos. As one example, antagonists to IGF1R and NRP1 are being developed as cancer therapeutics, and blocking their SNX5-mediated trafficking could provide an alternative approach for reducing receptor activity. The knowledge and the molecules we develop here will enable future studies to directly assess the effect of retromer targeting as a potential therapeutic strategy.

Chapter 6

Chapter 6

The GAP activity of SNX-RGS proteins

6.1 Introduction and significance

Regulator of G-protein signalling (RGS) proteins bind to activated Gα subunits proteins and serves as GTPase activating proteins (GAPs) by accelerating GTP hydrolysis and thus attenuating G-protein mediated signalling. To date, RGS proteins that are GAPs for G_i, $G\alpha_q$ and $G\alpha_{12/13}$ are the best studied, whereas details on RGS proteins that regulate $G\alpha_s$ are very limited (147-149).

Zheng *et al*., has previously identified the RGS domain of SNX13 as GAP protein for $G\alpha_s$, thus contributing to the regulation of cellular responses mediated by $G\alpha_s$ including adenyl cyclases and cAMP production (48). Recently SNX14 was found to bind and sequester $G\alpha_s$ even though it did not appear to regulate its GTPase activity (49). Neither the binding ability of SNX25 to $G\alpha_s$ nor its GAP activity has been studied so far.

The primary sequences of RGS domains and $G\alpha_s$ proteins may partly explain the specificity of the interaction between $G\alpha_s$ subunits and SNX-RGS proteins. The Asp229 of $G\alpha_s$ has been suggested to be the major barrier of $G\alpha_s$ interaction with other RGS proteins. Mutation of this residue to match the corresponding Ser206 of $G\alpha_i$ allowed the binding of RGS4 and RGS16 to the mutated $G\alpha_s$D229S. The crystal structure of the $G\alpha_{i1}$-RGS4 complex reveals that the Ser206 of $G\alpha_{i1}$ interacts with Glu126 and Asn128 of RGS4 to form a stable complex (150). Other amino acid residues occupy these positions in the RGS domains of SNX13, SNX14 and SNX25. This suggests that these nonconserved amino acid substitutions might contribute to the specificity of $G\alpha_s$-SNXRGS interaction.

In this chapter, I have expressed and purified the RGS domains from human SNX13, SNX14 and SNX25, as well as $G\alpha_s$. Using GST-pull down assays and isothermal titration calorimetry (ITC) experiments, I have systematically tested the binding affinity of the RGS domains of SNX13, SNX14 and SNX25 for $G\alpha_s$ in its inactive state ($G\alpha_s$-GDP bound), transition state mimic ($G\alpha_s$-GDP.AlF_4^- bound) and activated state ($G\alpha_s$-GPPNHP bound). I have shown that these RGS proteins bind to $G\alpha_s$ in both its active and inactive state, as opposed to previously published reports (48,49), although they display relatively low affinities for the inactive state. However, although the RGS proteins can form stable complexes with different $G\alpha_s$-nucleotide states, HPLC experiments indicate that the SNX-RGS domains do not appear to function as GTPase activating proteins for $G\alpha_s$. In this

chapter the first crystal structure of the SNX25 RGS domain begins to shed light on how these proteins might interact with and regulate $G\alpha_s$-mediated signaling. Future work will aim to determine the crystal structures of the RGS domains in complex with $G\alpha_s$ in different nucleotide-bound states to determine the precise residues defining specificity between these protein partners.

6.2 Materials and Methods

6.2.1 Construct and plasmid design

All bacterial expression constructs were synthesized and cloned into pGEX-4T-2 (RGS domains) or pET-28b(+) ($G\alpha_s$) by Genscript (USA). The constructs in pGEX-4T-2 are the RGS domains of SNX13, SNX14 and SNX25. The pGEX-4T-2 constructs are N-terminally GST-tagged with site- specific thrombin cleavage sequence located between the GST moiety and the protein of interest. The pET-28b(+) construct is N-terminally 6x His-tagged (H_6).

6.2.2 Recombinant protein production

The same protocol was used to express all the proteins mentioned under 6.2.1. The vectors encoding the genes of interest were transformed into competent *Escherichia coli* BL21 CODON-Plus cells. Cells were cultivated in Luria Bertani (LB) media at 37°C under agitation until the optical density at 600 nm (OD_{600}) reached 0.5-0.7 and expression was induced with 0.5 mM IPTG (except for $G\alpha_s$ where the OD_{600} was reached 0.8 and induced with 0.1 mM IPTG), and growth continued at 18°C for 16 h. The cells were harvested by centrifugation (4,000 x g, 15 min) and the resulting pellets resuspended in lysis buffer (50 mM Tris pH 8.0, 500 mM NaCl, 5% glycerol, 50 µg/mL benzamidine and 1 µg/mL DNAseI. The cell suspension was lysed using a cell disruptor at 27 Kpsi and cellular components were pelleted at 16000 RPM using JA 25.50 for 30 min. The supernatant was collected and the proteins were bound to an equilibrated glutathione sepharose 4B resin (GE Healthcare) or Talon metal affinity resin (Clontech) for 1 h at 4°C. The resins were washed with 10 column volumes of equilibration buffer (50 mM Tris pH 8.0, 500 mM NaCl, 5% glycerol buffer with and without 1 mM DTT for glutathione-S-transferase (GST) fusion RGS proteins and hexahistidine (H_6) $G\alpha_s$ respectively). The H_6-tagged $G\alpha_s$ protein was eluted in 5 fractions of 5 mL equilibration buffer containing 250 mM Imidazole. GST-tagged RGS proteins were cleaved on column by 100 units of thrombin overnight and were eluted the next day using the equilibration buffer. Both GST and H_6 tags were left uncleaved to perform invitro pull-down assays, whereas the GST tag was cleaved off for ITC and

crystallisation experiments. The cleaved proteins were then purified in gel filtration (GF) buffer (50 mM HEPES pH 8.0, 100 mM NaCl) by GF on a Superdex 200 16/60 size exclusion chromatography (SEC) column. These proteins were further dialysed into appropriate buffers according to the experiment performed.

6.2.3 Making $G\alpha_s$ – Guanine nucleotide complexes

100 µM $G\alpha_s$ protein was incubated with 1 mM GDP and GPPNHP in 20 mM HEPES pH 8.0, 200 mM NaCl, 1 mM DTT, 1 mM EDTA and with additional 10 µM GDP, 30 µM $AlCl_3$, and 10 mM NaF for the $G\alpha_s$ with $GDP{\cdot}AlF_4^-$ complex at room temperature. After an hour of incubation, 10 mM $MgCl_2$ was added to the mixture and was incubated for further 1 h. To remove excess nucleotides, these mixtures were then dialysed into buffers suitable for further experiments.

6.2.4 *In vitro* GST-pull downs

The RGS proteins were tested for interaction with $G\alpha_s$ in the presence of GDP, $GDP.ALF_4^-$ and GPPNHP using GST pull-down assays. GST-tagged RGS domains were dialysed into 50 mM HEPES pH 8.0, 200 mM NaCl, 1 mM DTT, 1 mM EDTA and 0.1% IGEPAL CA-630 and the $G\alpha_s$–nucleotide complexes were dialysed into 50 mM HEPES pH 8.0, 200 mM NaCl, 1 mM DTT, 1 mM EDTA, 10 mM $MgCl_2$ and 0.1% IGEPAL CA-630 for GDP and GPPNHP bound states, whereas the $G\alpha_s$–$GDP.ALF_4^-$ complex was dialysed into 50 mM HEPES pH 8.0, 200 mM NaCl, 1 mM DTT, 1 mM EDTA, 10 mM $MgCl_2$, 10 µM GDP, 30 µM $AlCl_3$, 10 mM NaF and 0.1% IGEPAL CA-630. 1.25 nmol of RGS proteins were mixed with 2.5 nmol H_6- $G\alpha_s$-nucleotide complexes for 1 h at 4°C, After 1 h incubation, these mixtures are bound to 250 µL of pre-equilibrated glutathione sepharose resin in respective pull-down buffers and were incubated for further 1 h. The resin was washed 4 times with 1 ml of appropriate pull down buffer. After the last spin, the supernatant was pipetted off, and the beads were spun full speed for 30 extra secs. Equal volume of SDS-sample buffer was added and SDS-PAGE analysis was performed. Western blot analysis was done using a nitrocellulose membrane and the iBlot semi-dry transfer system (Life technologies). Detection was performed using a primary mouse anti-$(His)_6$ and secondary goat anti-mouse ECL-coupled antibodies and a photographic film with the BioRad ECL detection system.

6.2.5 Isothermal titration calorimetry (ITC)

Isothermal titration calorimetry (ITC) was performed at 10°C using a MicroCal PEAQ-ITC calorimeter. Both RGS and $G\alpha_s$–nucleotide protein samples were dialysed against 50 mM

HEPES pH 8.0, 100 mM NaCl, 10 mM $MgCl_2$, and 1 mM EDTA for measurement of the GDP and GPPNHP states, and with an additional 10 µM GDP, 30 µM $AlCl_3$, and 10 mM NaF for the $G\alpha_s$-GDP·AlF_4^- complex. Contents of the sample cell were stirred continuously at 750 rpm during the experiment. 13 injections were involved in an experiment and 50 µM of $G\alpha_s$ were titrated into 500 µM of RGS proteins. The binding data was processed using MicroCal PEAQ-ITC analysis software with a single site-binding model to determine the stoichiometry (n), the binding constant (K_d), the estimated heat of binding (ΔH), the Gibbs free energy (ΔG) and the binding entropy (ΔS).

6.2.6 GTPase activation assay experiments using HPLC

The nucleotide analyses were performed using C-18 reversed phase column (Aeris peptide from Phenomenex; 250 mm x 4.6 mm) filled with 3.6 µm (particle size) silica. The column was equilibrated with buffer containing 100 mM potassium phosphate buffer (pH 6.5), 10 mM tetrabutylammonium bromide and 7.5% acetonitrile, and isocratic elution was performed at ambient temperature with a flow rate of 0.750 mL/min. The nucleotides were detected at 254 nm absorbance. In this system, the order of elution of guanine nucleotides was GDP and GTP (retention times of ~ 7 and 12 min, respectively). 8 µM $G\alpha_s$ protein was mixed with 80 µM RGS protein in GAP buffer (30 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 3 mM DTT, 10 mM potassium phosphate buffer pH 7.4) and were pre-incubated at room temperature. 100 µM GTP solution was added to the premixed GTPase solution. 30 µL samples were collected at different time points (0, 1, 2, 5, 10, 15, 20 and 30 mins) and were snap frozen in liquid nitrogen. Each sample was boiled for 3 min at 100°C and centrifuged at 13000g for 10 min to remove the denatured protein. An aliquot of the supernatant (20 µL) was injected into the HPLC column. The column was calibrated with 25, 50, 100, and 200 µM solutions of the different guanine nucleotides (GDP and GTP). The relative amount of GTP was calculated using the GTP and GDP peak areas measured using the HPLC integrator and were introduced into the following equation:

$$Relative\ amount\ of\ GTP = \frac{GTP}{GTP + GDP}$$

The GTP content (Y-axis) was then plotted against the time of incubation (X-axis) using a single exponential function.

6.2.7 Protein crystallisation, data collection and structural determination

Initial crystallisation screens in small volumes were set up using four 96-well crystallisation screens for each protein sample. 8 mg/ml of SNX13, SNX14 and SNX25 RGS domains were mixed and incubated on ice with $G\alpha_s$-GDP, $G\alpha_s$-GDP.ALF_4^-, and $G\alpha_s$-GPPNHP in a

1:1 stoichiometry. RGS domains in their apo state were also set up for crystallisation. 200 nL 1:1 hanging drops of proteins were dispensed using the Mosquito robot (TTP labtech) in the UQ ROCX facility. After setting up these initial screens, the trays were incubated at 20°C in a Rockimager storage hotel (Formulatrix), where the drops were imaged at different time points. Data were collected at the Australian Synchrotron MX2 Beamlines, integrated with iMOSFLM (93), and scaled with AIMLESS (94) in the CCP4 suite (95). The structure of the SNX25 RGS domain was solved by molecular replacement with PHASER using the RGS1 structure as input (PDB ID 2BV1) (96). The resulting model was rebuilt with COOT (114), followed by repeated rounds of refinement with PHENIX (98). All structural figures were generated using PyMOL (DeLano scientific).

6.6 Results

6.6.1 Recombinant protein expression and purification

Purifications of SNX13-RGS, SNX14-RGS and SNX25-RGS were performed using a two-step protocol; affinity chromatography with glutathione sepharose followed by gel filtration using Superdex 75 (16/600) column. The proteins were eluted at the expected molecular weight. Samples from every step of purification were collected for SDS-PAGE and were analysed for the presence of successfully purified proteins of interest (**Fig. 55A, 55B and 55C**). Fractions were found to be relatively pure and dimeric. Upon addition of DTT, the protein seems to be monomerised (**Fig. 55A-C**, Lane 3), which lead us to believe the dimer was formed through disulphide bonds. The fractions were pooled together for further experiments. After the gel filtration, the yield of SNX13-RGS, SNX14-RGS and SNX25-RGS were around 15 mg/mL, 18 mg/mL and 10 mg/mL respectively.

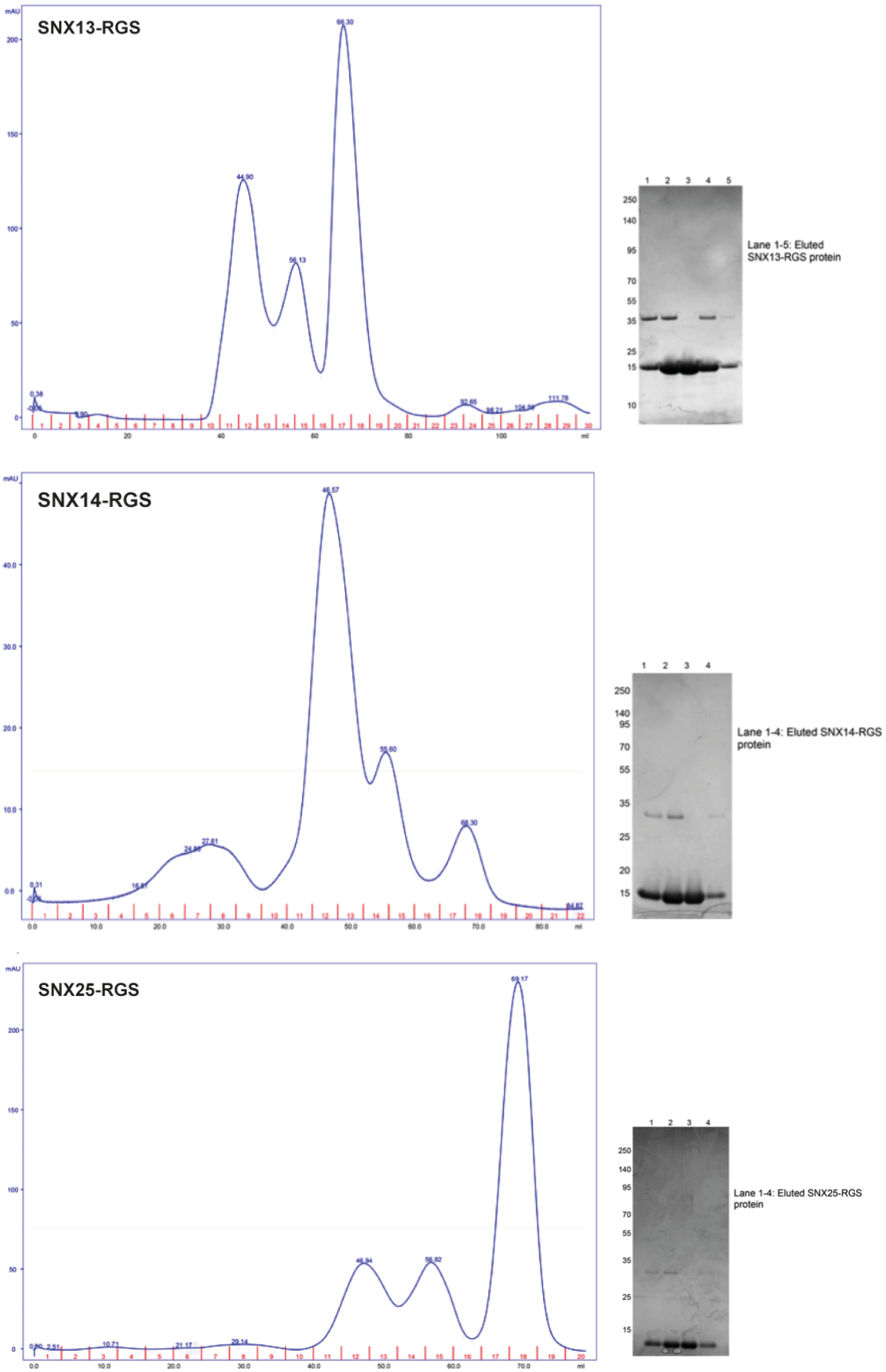


***Figure 55**. FPLC chromatography profile and SDS-PAGE of untagged RGS domains of (A) SNX13 (17.3 kDa), (B) SNX14 (16.5 kDa) and (C) SNX25 (14 kDa). The fractions from the major peak between 60-80 mls correspond to the single bands on the gel.*

6.6.2 GST-pull down experiments to verify the interaction between SNX-RGS proteins and $G\alpha_s$-nucleotide complexes

The GST pull-down assay was performed to obtain a qualitative understanding about the interaction between RGS proteins and the $G\alpha_s$ in its inactive (GDP-bound form), active (GPPNHP-bound form) and transition state conformations (GDP.AlF_4^- bound form). As opposed to previous observations (48), the RGS domains of SNX13 and SNX14 were found to interact with the inactive form of $G\alpha_s$, whereas SNX25 showed no binding (**Fig. 56**).

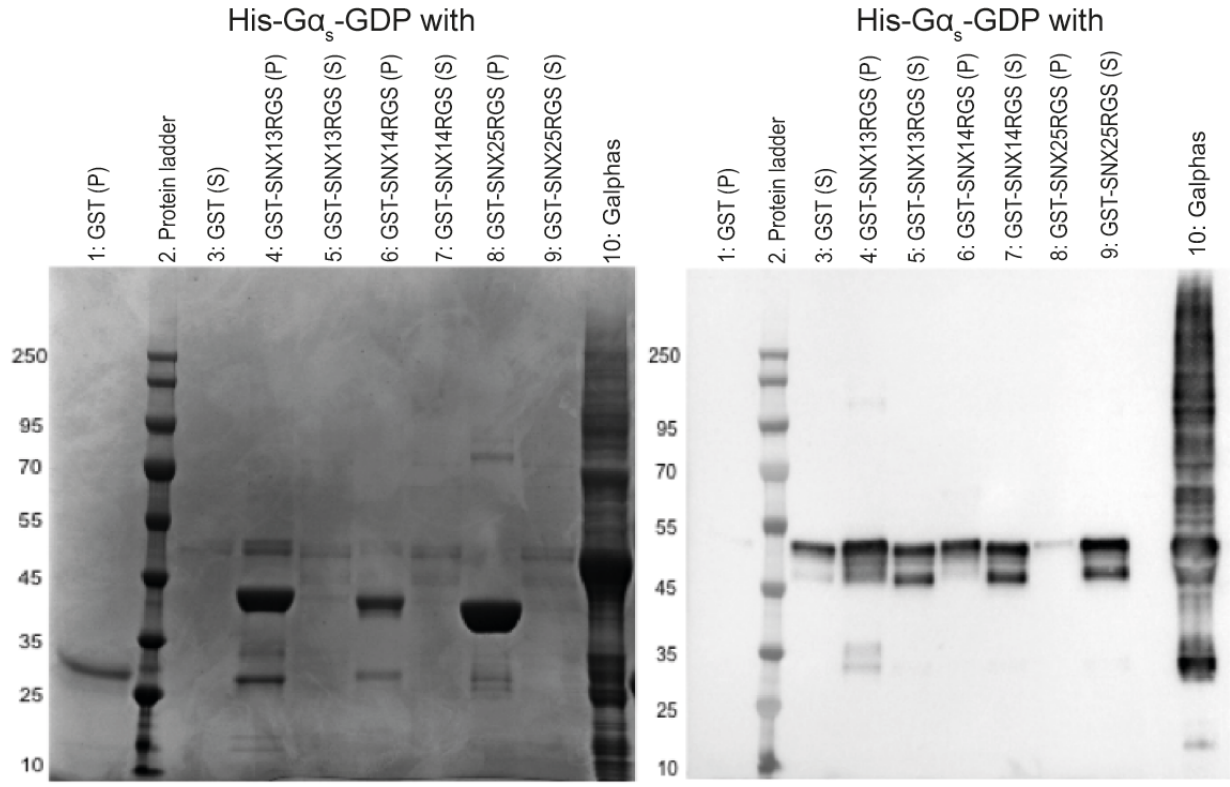


Figure 56. *GST pull down experiment of GST-RGS proteins vs His-$G\alpha_s$-GDP.* ***(Left)*** *SDS-PAGE gel and* ***(Right)*** *western blot containing the Pellet (P) or Resin and supernatant (S) fractions from the experiment. Experiment between GST and His-$G\alpha_s$-GDP was performed as the negative control. GST = 26 kDa, His-$G\alpha_s$-GDP=46 kDa, GST-SNX13RGS=43 kDa, GST-SNX14RGS=42.5 kDa, GST-SNX25RGS=40 kDa and His-$G\alpha_s$ protein=46 kDa.*

SNX13-RGS protein showed interaction with both the active and transition forms of $G\alpha_s$, whereas SNX14-RGS showed preference for the GDP. AlF_4^- bound state of $G\alpha_s$ over the GPPNHP bound state (**Fig. 57**). SNX25 showed interaction with both GDP.AlF_4^- and GPPNHP bound states of $G\alpha_s$. In contrast, in the control pull-down assay, none of the combinations of the GST-alone protein with the His-tagged $G\alpha_s$-nucleotide proteins interacted in the analyses (**Figs. 56 and 57**).

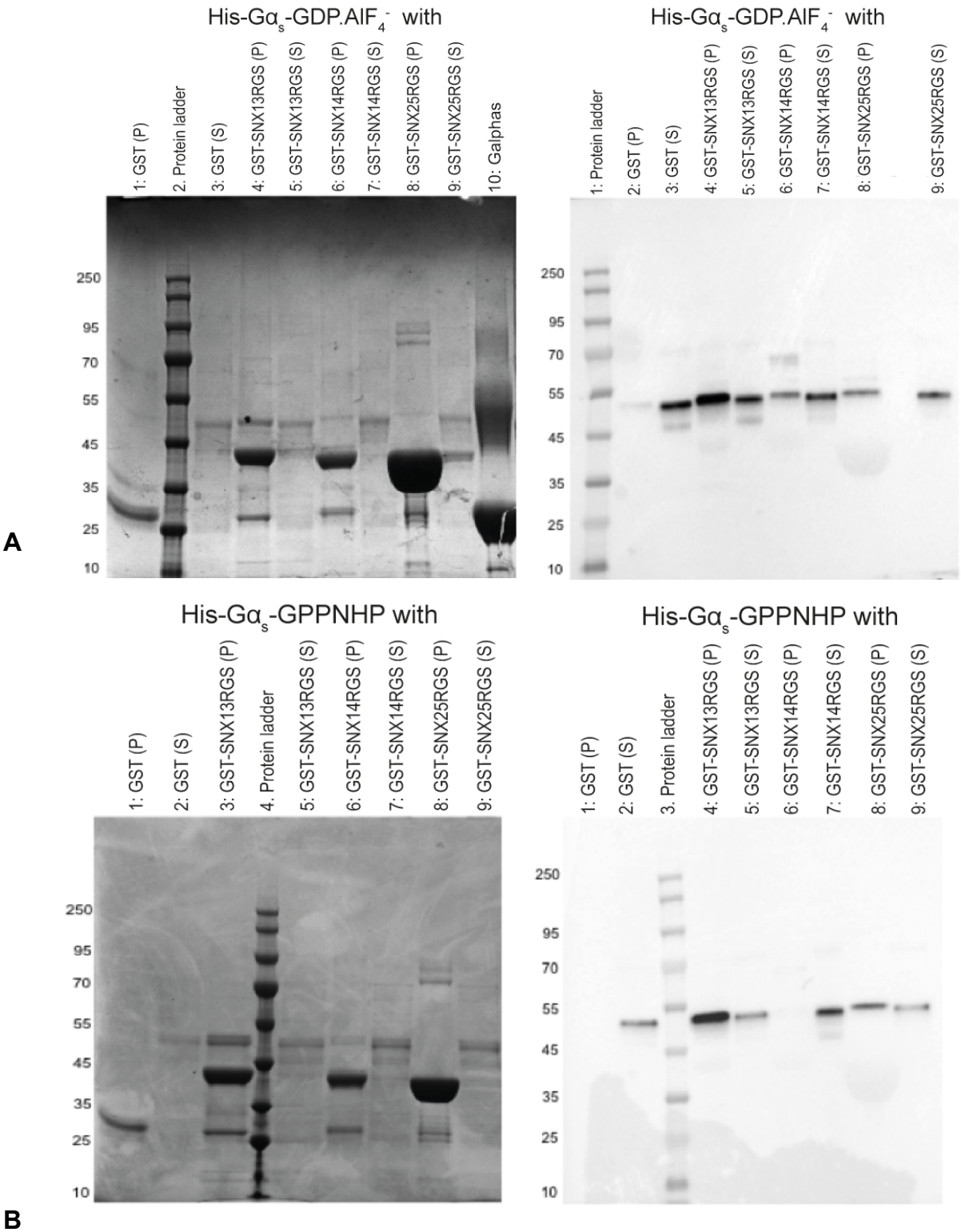


***Figure 57**. GST pull down experiments of **(A)** GST-RGS proteins vs His-Gα$_s$-GDP.AlF$_4^-$ and **(B)** GST-RGS proteins vs His-Gα$_s$-GPPNHP. **(Left)** SDS-PAGE gel and **(Right)** western blot containing the Pellet (P) or Resin and supernatant (S) fractions from the experiments. Experiment between GST vs His-Gα$_s$-GDP.AlF$_4^-$ and His-Gα$_s$-GPPNHP were performed as the negative control. GST = 26 kDa, His-Gα$_s$-GDP.AlF$_4^-$ and His-Gα$_s$-GPPNHP =46 kDa, GST-SNX13RGS=43 kDa, GST-SNX14RGS=42.5 kDa, GST-SNX25RGS=40 kDa and His-Gα$_s$ protein=46 kDa.*

6.6.3 RGS domains of SNX-RGS proteins bind to both inactive and active forms of $G\alpha_s$

Previous studies report that most RGS domains typically bind to the transition state conformation ($GDP{\cdot}AlF_4^-$) of Gα proteins with the highest affinity (151-153). To measure the binding affinities of the SNX-RGS domains for different $G\alpha_s$ nucleotide state I used ITC. The dissociation constants for binding of SNX13RGS with the $GDP{\cdot}AlF_4$ and GPPNHP forms of $G\alpha_s$ were 6 μM and 14.3 μM respectively (**Fig. 58**).

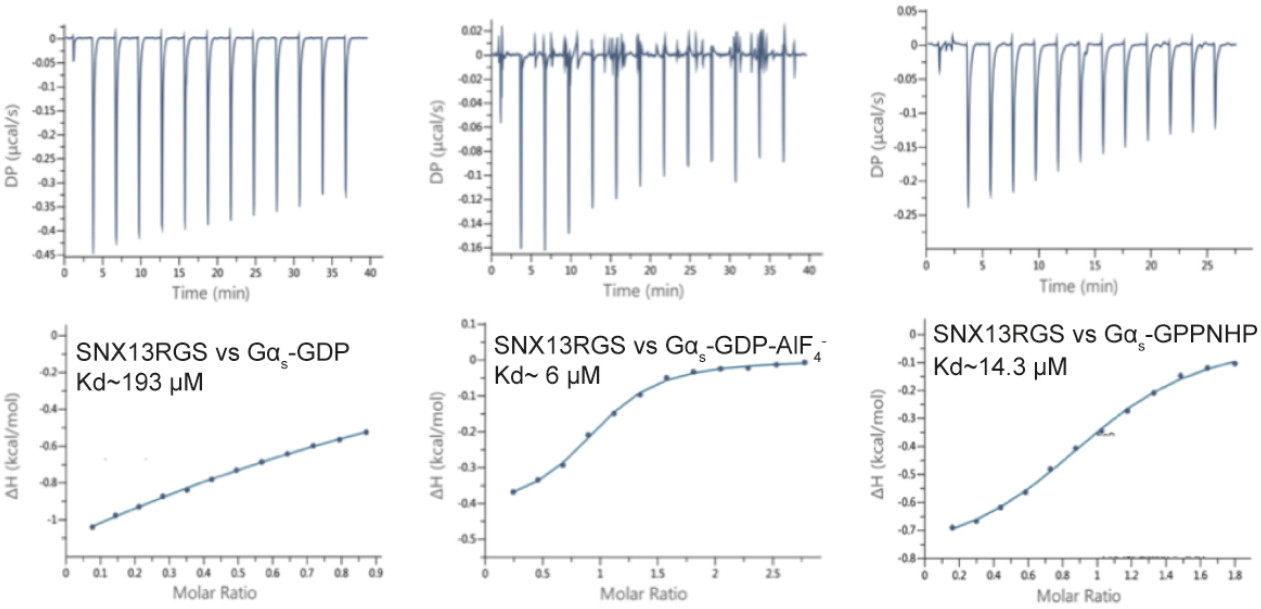


Figure 58. *ITC profiles for the binding of the RGS domain of SNX13 vs $G\alpha_s{\cdot}GDP$* ***(left)****, $G\alpha_s{\cdot}GDP{\cdot}AlF_4^-$* ***(middle)****, and $G\alpha_s{\cdot}GPPNHP$* ***(right)****. Thermodynamic parameters are presented in Table* ***8*** *at the end of this chapter.*

SNX14RGS bound preferentially to $GDP{\cdot}AlF_4$ form of $G\alpha_s$ with a K_d of 12.5 μM and showed no binding to $G\alpha_s$-GPPNHP (**Fig. 59 top**). In contrast, SNX25RGS was found by ITC to bind with highest affinity (K_d ~ 40 μM) to the GPPNHP bound form of $G\alpha_s$ and with somewhat lower affinity (K_d ~ 72.4 μM) to the $G\alpha_s$-$GDP{\cdot}AlF_4^-$ complex (**Fig. 59 bottom)**. SNX13 and SNX14 RGS domains have lower, but still substantial affinity for the $G\alpha_s$-GDP complex (K_d ~ 193 μM and 66 μM respectively) (**Fig. 58-59**) (**Table 8**).

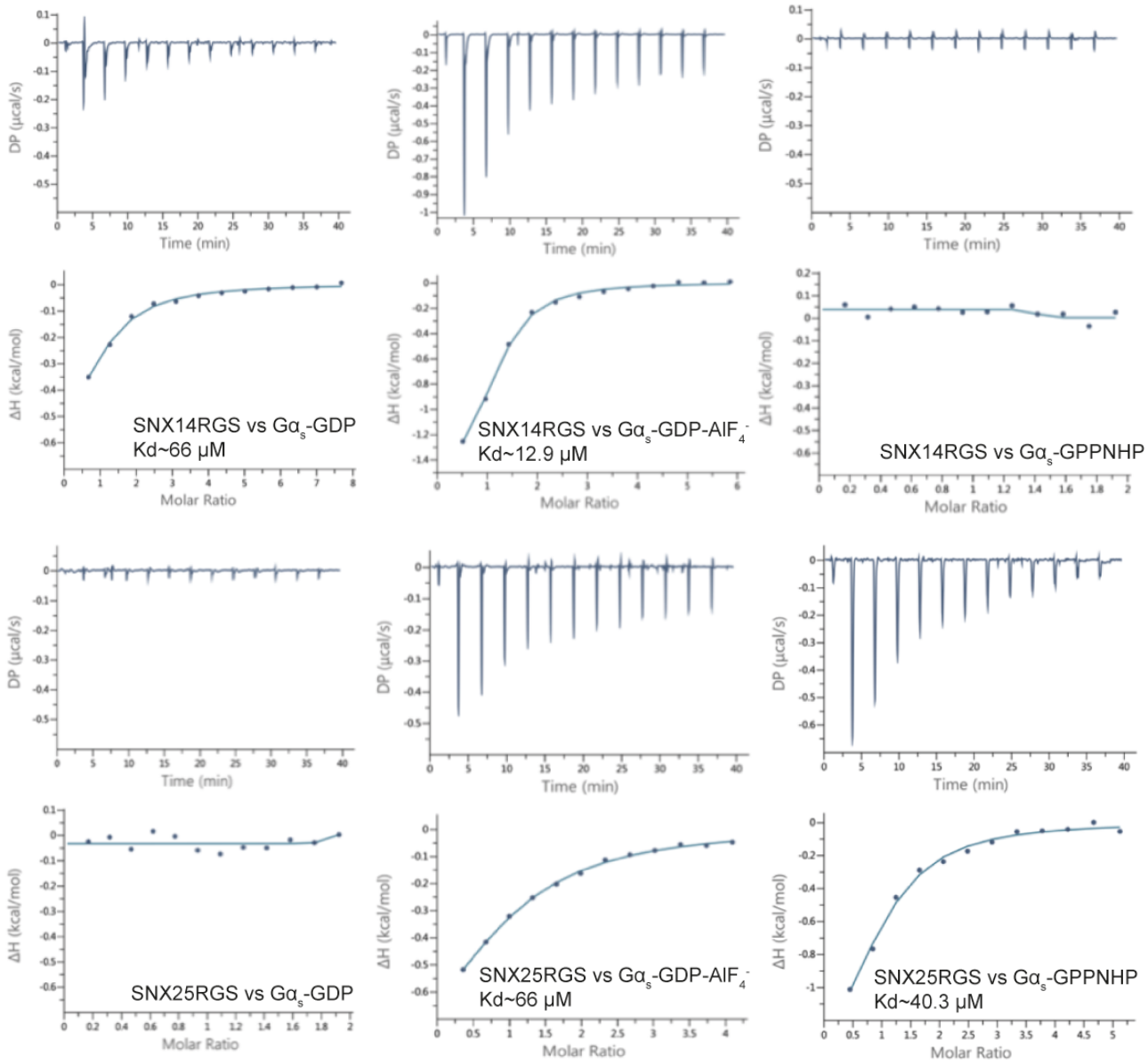


Figure 59. *ITC profiles for the binding of the RGS domain of* ***(Top)*** *SNX14 and* ***(Bottom)*** *SNX25 vs* $G\alpha_s \cdot GDP$ ***(left)***, $G\alpha_s \cdot GDP \cdot AlF_4^-$ ***(middle)***, *and* $G\alpha_s \cdot GPPNHP$ ***(right)***. *Thermodynamic parameters are presented in Table* ***8*** *at the end of this chapter.*

6.6.4 GAP activity

Previous studies report SNX13RGS possess GAP activity for $G\alpha_s$, whereas SNX14RGS protein do not. Whether SNX25RGS possess GAP activity is unknown. To understand whether the RGS domains of SNX13, SNX14 and SNX25 possess GAP activity for $G\alpha_s$, HPLC assay was used. The preliminary results demonstrated no conversion of GTP to GDP by the SNX-RGS proteins (**Fig. 60**). To confirm the assay protocol, positive control

experiments are currently being performed. RGS16 is a GAP protein of $G\alpha_{iq}$ subunit and these constructs will be used to perform positive control assays.

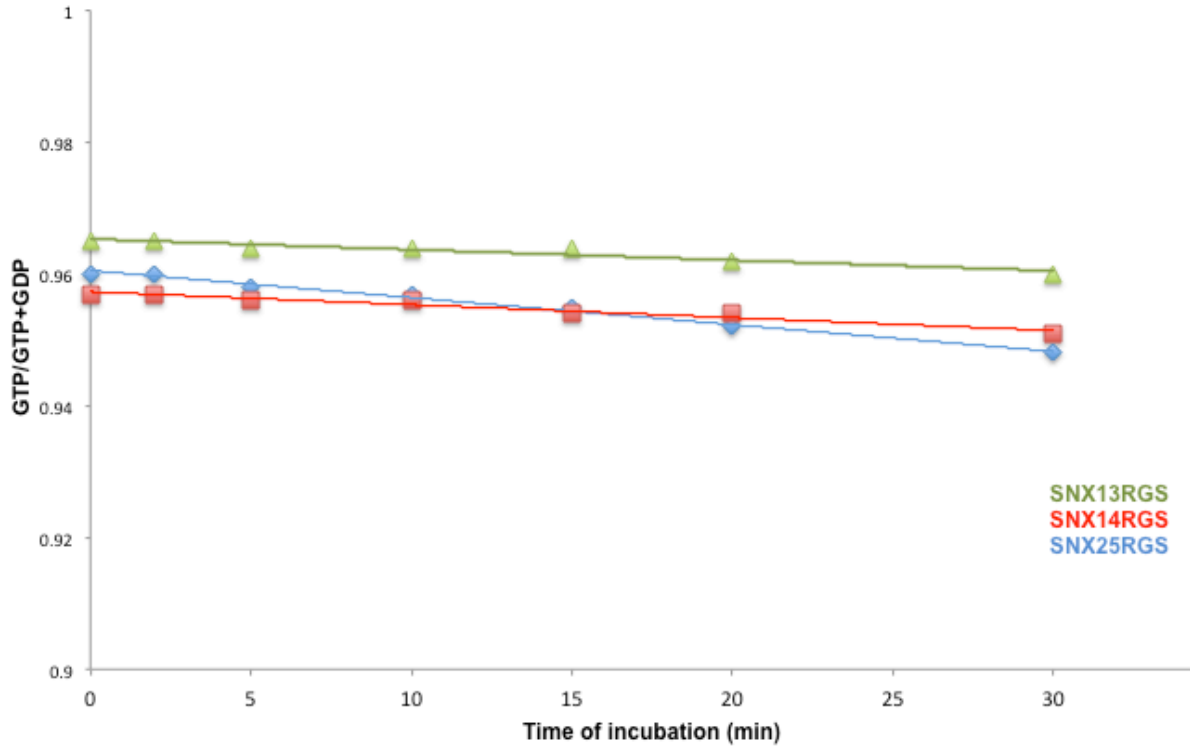


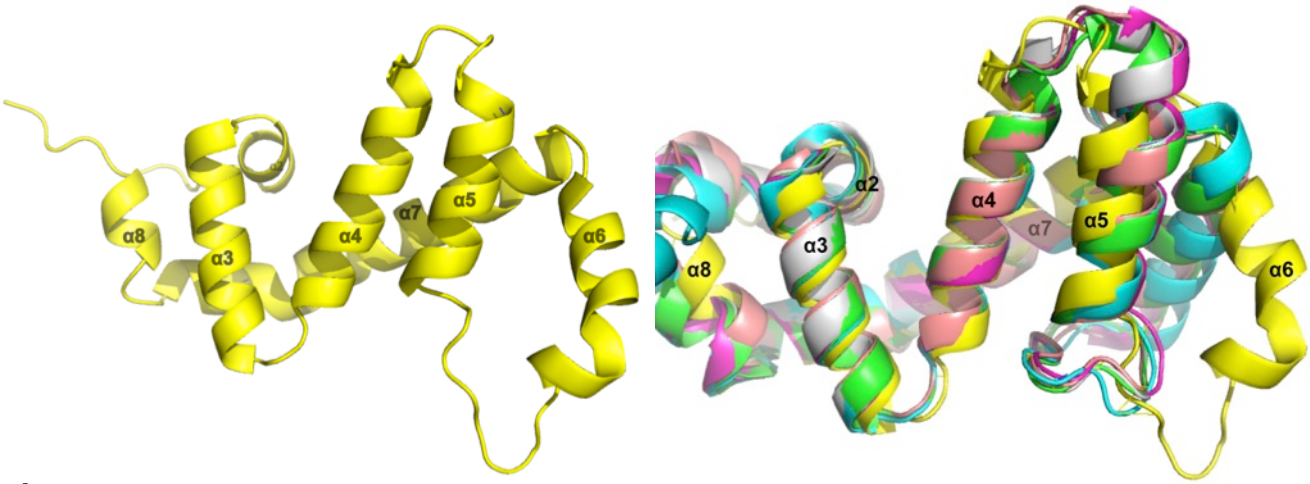
Figure 60. *HPLC-based analysis of the SNX-RGS stimulated GAP activity.*

6.6.5 Crystal structure of apo-SNX25 RGS

I next used X-ray crystallography to determine the crystal structure of the apo RGS domain of SNX25 as a path to understanding its complex structure with $G\alpha_s$. After setting up 96-well screens I identified many crystal hits. Around six different conditions with crystal hits were immediately set up in a 24-well sitting drop plate to optimize crystal size for data collection. Almost all conditions gave larger well-formed crystals. An apo SNX25RGS crystal (**Fig. 61**) (reservoir condition: 30% PEG 4000, 0.2 M Sodium acetate, 0.1 M Tris pH 8.5) was sent to the Australian Synchrotron and the structure was solved to a resolution of 2.4 Å using molecular replacement (**Table 9**).

***Figure 61**. Crystals of apo SNX25RGS in a 24-well plate. The reservoir condition is 30% PEG 4000, 0.2 M Sodium acetate, 0.1 M Tris pH 8.5.*

The majority of apo-RGS domain structures conform to the typical structural fold seen in RGS4, which is a nine-helix bundle with two lobes formed by the I, II, III, VIII, and IX helices and the IV, V, VI, and VII helices, respectively. The SNX25 RGS protein also conforms to the canonical folds as seen in **Fig. 62A**. When superimposed with some of the other known structures of RGS, the α6 helix seemed to have a tilt, which also causes a conformational change in the α5- α6 loop as well (**Fig. 62B**). There is a cysteine residue where the tilt is seen in α6 helix, which makes a disulphide bond with a neighbouring protein in the crystal lattice. This needs to be further examined to determine if it is influencing the unusual structure of the SNX25 RGS domain, especially as the α5- α6 helical area of RGS proteins is known to interact with Gα subunits.



A **B**

***Figure 62.** (A) Apo SNX25 RGS structure with the alpha-helical folds as expected; (B) Superimposition of apo SNX25 RGS (Yellow) with other RGS such as RGS1 (green, PDB code: 2BV1), RGS2 (Cyan, PDB Code: 2AF0), RGS4 (Peach, PDB code: 1AGR), RGS8 (Magenta, PDB code: 2IHD) and RGS10 (Grey, PDB code: 2IHB).*

Table 9. Summary of crystallographic structure determination statistics

Data collection	
Wavelength (Å)	1.0064
Space group	P121
Cell dimensions	
a, *b*, *c* (Å)	48.61, 57.12, 66.98
a, b, g (°)	90, 109.49, 90
Resolution (Å)	63.2-2.53 (2.69-2.53)
R_{merge}	0.025 (0.373)
R_{meas}	0.079 (0.508)
R_{pim}	0.025 (0.343)
<*I*> / s*I*	14.5 (3.2)
Total number reflections	42496 (4932)
Total unique reflections	11564 (1381)
Completeness (%)	99.6 (98.3)
Multiplicity	3.7 (3.6)
Half-set correlation (CC(1/2))	0.996 (0.864)
Refinement	
Resolution (Å)	45.8- 2.5 (2.65-2.53)
No. reflections/No. R_{free}	10881/ 1107
R_{work}/R_{free}	0.186/0.235(0.193/ 0.231)
No. atoms	
Protein	2083
Solvent	20
Average *B*-factor (Å^2)	33.5
R.m.s deviations	
Bond lengths (Å)	0.016
Bond angles (°)	1.931

6.6.6 Crystallisation trials of RGS- $G\alpha_s$ complexes

Out of different initial screens that I tried in 96-well trays, several conditions produced microcrystals of SNX13RGS-$G\alpha_s$.GDP.AlF_4^-, SNX14RGS-$G\alpha_s$.GDP.AlF_4^-, and

SNX25RGS-Gα$_s$.GPPNHP (**Fig. 63**). These conditions are currently being optimised to obtain macrocrystals good enough for a high-resolution structural analysis.

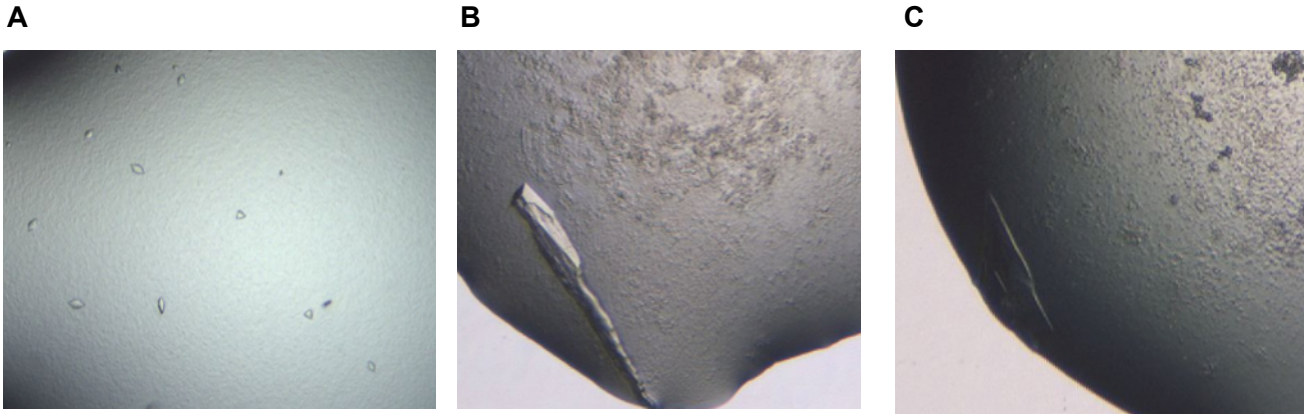


Figure 63. *96-well plate crystals of SNX13RGS-Gα$_s$.GDP.AlF$_4^-$ (30% PEG 400, 0.1 M CAPS pH 10.5, 0.5 M $(NH_4)_2SO_4$, 10%(v/v) glycerol; SNX14RGS-Gα$_s$.GDP.AlF$_4^-$,(30% PEG 200, 0.1 M TRIS pH 8.5, 0.2 M $(NH_4)_2HPO_4$ and SNX25RGS-Gα$_s$.GPPNHP (0.05 M Tris HCl pH 8.5, 22%(w/v) PEG 4000).*

6.7 Discussion

The drug selectivity of G-protein-coupled receptors has been therapeutically used to target various cellular processes (154-156). The selectivity of G-proteins for different RGS proteins is now also being considered as a pharmacological target for the regulation of G-protein-regulated signaling (157,158). So far, the activity of different G-proteins in specific signaling pathways has been studied in great detail, but the role of RGS proteins exhibiting selectivity and GAP activity for these G-proteins is largely unexplored, particularly for the pharmacologically important Gα$_s$. Most of the first known RGS family members were shown to have GAP activity towards Gα proteins (42,159,160). However, although RGS proteins bind to Gα subunits, not all RGS domains function as GAPs for GTP hydrolysis.

The results of this chapter demonstrate the ability of the RGS domains of SNX-RGS proteins (SNX13, SNX14 and SNX25) to bind to Gα$_s$ using *in vitro* GST-pull down and ITC experiments. Further, we report that there are differences in the relative affinities of the SNX-RGS proteins for the different Gα$_s$- nucleotide complexes. In general there is a preference for either the transition state conformation or the active state of the Gα$_s$ protein, although there is also a weaker interaction between SNX13 and SNX14 with the inactive state of the molecule. Finally, my studies demonstrated that even though all these SNX-RGS proteins successfully interact with Gα$_s$, none of them were found to possess GAP activity. Positive control assays will be performed inorder to confirm this assay protocol

and thus validate my findings.

In contrast to the previous reports by Zheng *et al*. (48), the RGS domains of SNX13 and SNX14 interacted with both inactive and active form of $G\alpha_s$ (although more strongly with the latter), whereas SNX25 showed no binding to the inactive form. SNX13 RGS protein showed interaction with the active and transition forms of $G\alpha_s$, whereas SNX14-RGS showed preference for the transition state over the active state. On the contrary, SNX25 showed preferential interaction with the active form over the transition state of $G\alpha_s$. While this confirms that each of these proteins possess the amino acids necessary for the interaction with $G\alpha_s$, the preference of RGS domains for one $G\alpha_s$–nucleotide bound state over the other suggests that the binding sites of the RGS proteins adopting different favourable conformations to bind the preferred $G\alpha_s$ state.

Although GAPs have been described for most G-proteins, a GAP for $G\alpha_s$ remained elusive until the original studies of SNX13-RGS appeared to answer this question, identifying a potential $G\alpha_s$-GAP activity (48). But no follow-up studies about this GAP activity have been reported since. Ha *et al*. recently reported that SNX14-RGS, even though it successfully binds and sequesters $G\alpha_s$, does not appear to regulate its GTPase activity (49). A potential flaw in this study however, is that no comparison was made with SNX13 (or SNX25), and in addition no positive control was used to confirm that the assay was working robustly. No studies have been reported on the GAP activity of SNX25-RGS. In this work, I have so far not been able to measure GAP activity for any of the SNX RGS domains, supporting the results of Ha *et al*. on SNX14, but in contrast to the reports by Zheng *et al*. on SNX13. The next steps in order to confirm this assay using a positive control are currently underway. Even though all three SNX-RGS proteins possess the required residues for the stable $G\alpha_s$ binding, my initial results did not demonstrate GAP acitivity. If this is the case, then the SNX-RGS proteins may still act as negative regulators of $G\alpha_s$ signaling by the mechanism proposed by Ha *et al*., whereby they will sequester the G-protein and prevent its interactions with downstream effectors. Alternatively, the SNX-RGS proteins may not be regulators of $G\alpha_s$, but may in themselves be specific effectors of these signaling molecules.

RGS proteins have previously been shown to be regulated either by phosphorylation, or by palmitoylation on cysteine residues. For example, RGS2 showed decreased GAP activity following protein-kinase C (PKC)-mediated phosphorylation (161), whereas RGS-GAIP showed increased GAP activity after ERK-mediated phosphorylation (162). RGS4 and RGS10 demonstrated decreased GAP activity through the palmitoylation of Cys^{95} and Cys^{66} residues respectively (163). Even though my studies demonstrate the

RGS domains of SNX-RGS do not possess GAP activity for $G\alpha_s$, both SNX13 and SNX14 proteins are previously reported to regulate cAMP signaling (48,49). As mentioned above, Ha *et al.* suggested the SNX14 RGS domain can regulate cAMP signalling without GTPase activity by directly sequestering $G\alpha_s$ and thus inhibiting downstream cAMP production, essentially resulting in a similar functional outcome to activating GTP hydrolysis. In addition, Protein Kinase A (PKA)- mediated phosphorylation of the SNX14-RGS domain at Ser382/Ser388 considerably decreased its affinity for $G\alpha_s$ thus reducing its inhibition of $G\alpha_s$, which in turn promotes downstream cAMP signalling pathways (49). When compared to other known structures of RGS, the structural data of apo SNX25-RGS shows an unsual tilt where there is a cysteine (Cys^{76}) residue on the α6 helix. Whether SNX13-RGS and SNX25-RGS share similar pathways of feedback phosphorylation or palmitoylation for the regulation of cAMP signalling are yet to be explored. These results contribute to the suggestion that $G\alpha_s$ proteins may differ from other G-proteins with regard to its regulation by GAPs.

The next intriguing result from this chapter was the RGS domains of SNX13 and SNX14 (193 μM and 66 μM respectively) formed a relatively low affinity interaction with the $G\alpha_s$ in its GDP-bound state. While the affinity of SNX-RGS for the $G\alpha_s$-GDP was considerably lower than for the stable and transition state of $G\alpha_s$, the interaction is significant and leads us to believe that the inactive $G\alpha_s$ subunit can potentially be functionally important in its interaction with SNX-RGS proteins. Though with lower affinity, $G\alpha_s$-GDP had been previously reported to stimulate the activity of adenylyl cyclase (164). This might suggest that GTP hydrolysis by Gα alone is not enough to fully terminate signaling and further sequestration of $G\alpha_s$-GDP by other proteins such as SNX-RGS may play a key role. Other RGS proteins have been shown to possess guanine nucleotide dissociation inhibitor (GDI) activity due to their ability to bind to the inactive state of Gα subunits (165,166). Whether SNX-RGS proteins possess GDI activity to promote the suppression of activation of $G\alpha_s$ are yet to be explored.

In the future, the crystal structures of the RGS domains in complex with $G\alpha_s$ will enable us to understand the residues providing specificity between these partners and how they achieve preference for the different $G\alpha_s$-nucleotide states. The longer-term goal will be to properly define the physiological purpose of SNX-RGS proteins in regulating $G\alpha_s$ mediated cAMP signaling. As discussed in previous chapters, SNX-RGS proteins have been implicated as key players in ER-endosomal membrane tethering that controls trafficking, signaling and exchange of lipids between the ER, LDs and endosomal compartments. Although the role in GPCRs in ER and LD function is complex and poorly

understood, it is interesting to note that the activation of Gαs-coupled β-adrenergic receptors results in phosphorylation of hormone sensitive lipase (HSL), rapid recruitment of HSL to LDs and subsequent production of free-fatty acids from stored triacylglycerol (167). This suggests one possible connection between the SNX-RGS proteins and GPCR mediated cellular signalling pathways, and these integrated processes of inter-organelle exchange and protein and lipid homeostasis are widely emerging as potential therapeutic targets in a number of neurodegenerative diseases.

Table 8. Thermodynamic parameters of the RGS proteins- $G\alpha_s$ interactions

Sample cell	Titrant	ΔH (kcal/mol)	TΔS (kcal/mol)	ΔG (kcal/mol)	K_d (uM)	N
SNX13 RGS	$G\alpha_s$·GDP	-2.01	-2.85	-4.85	193	1.05
	$G\alpha_s$·GDP.AlF_4^-	-0.389	-6.42	-6.81	6	1.09
	$G\alpha_s$·GPPNHP	-0.815	-5.51	-6.32	14.3	1.01
SNX14 RGS	$G\alpha_s$·GDP	-0.704	-4.75	-5.45	66	1.18
	$G\alpha_s$·GDP. AlF_4^-	-1.66	-4.72	-6.38	12.9	1.01
	$G\alpha_s$·GPPNHP	/	/	/	/	/
SNX25 RGS	$G\alpha_s$·GDP	/	/	/	/	/
	$G\alpha_s$·GDP. AlF_4^-	-1.02	-4.38	-5.40	72.4	1.09
	$G\alpha_s$·GPPNHP	-1.55	-4.19	-5.78	40.3	1.16
ΔH – enthalpy, T- absolute temperature, ΔS –change in entropy, ΔG – Gibbs free energy, K_d – dissociate constant, N – stoichiometry						

Chapter 7

Chapter 7

Towards the structure and function of the unique PXA domain from the SNX-RGS protein family

7.1 Introduction and significance

The SNX-RGS subfamily of proteins consists of four human homologues SNX14, SNX13, SNX19 and SNX25, and a single homologue in *Saccharomyces cerevisiae* called Mdm1p. This subfamily is unique amongst the other SNX proteins because they possess transmembrane domains (**Chapter 1, Fig. 9**). Besides the two N-terminal transmembrane helices this subfamily also contain an N-terminal PX-associated domain (PXA), a regulator of G-protein coupled receptor (RGS) domain (present in all proteins except SNX19), a central PX domain, and a C-terminal PX-associated domain (PXC) (8,38).

Recent studies have demonstrated the involvement of SNX14 in human neurological disease. SNX14 mutations were found in cerebellar ataxia, which is a condition caused by inflammation in the cerebellum of Purkinje cells, resulting in defective control over muscle movements. At the cellular level the major observation in both patients and morpholino depleted zebrafish was an accumulation of autophagic structures containing undigested material, suggesting a deficiency in lysosomal maturation and normal cellular degradation (54,55). Henne *et al*. (2015) reported perhaps the most insightful mechanistic studies on the cellular role of SNX-RGS proteins, demonstrating that the yeast homologue Mdm1p is an ER-vacuole tethering protein. Overexpression of Mdm1p led to enlarged nuclear ER-vacuole junctions (NVJs), which suggests SNX14-related proteins can act as interorganelle tethering proteins by anchoring them to ER via their two N-terminal transmembrane domains and by binding to the phospholipids on the endosomal membranes through the PX domain. MDM1 deletions based on the disease-causing truncations in human SNX14 in turn lead to unsuccessful localisation to the NVJs (59). In a more recent study, Henne et al., found that Mdm1p localises to ER-vacuole junctions and regulates lipid droplet budding during nutrient starvation (60). They are also found to interact with fatty-acyl CoA synthases including Faa1 (60), potentially through the PXA domain (Mike Henne, *personal communication*). Thin layer chromatography of purified Mdm1p PXA domain revealed the presence of co-purified long chain free fatty acids and the lipids potentially seen on the lipid droplet monolayer (Mike Henne, *personal communication*). The accumulation of lipid droplets at the ER-vacuole interface could be a

strategy used by the cell to regulate the formation of lipid droplets where they maybe easily transferred to the vacuole and eventually digested through lipophagy (168).

These studies lead us to believe that SNX-RGS proteins can encourage membrane contacts at the ER-endosome sites that regulate normal lipid metabolism and transport. The disease causing truncations in SNX14 can lead to defective ER, LD and endosomal communication, which could subsequently lead to defective lipophagy or autophagy, which leads to neurological disease due to disrupted neuronal homeostasis. Although SNX-RGS protein family members are highly conserved and are involved in such important and diverse roles there is still no proper understanding of the structural or functional details of these proteins. Our lab recently published crystal structures of the PX domains from SNX14 and SNX19 (38), and in the previous chapter I described my structure and characterisation of the RGS domains from this protein family. In this chapter I focussed on the PXA and PXC domains of the SNX-RGS proteins as these domains are completely structurally uncharacterised, and any new structural information should provide significant new insights into the molecular mechanisms by which these domain bind to lipids and proteins (50,51), and how they function in regulation of lipid droplet biogenesis at the ER-LD contact site.

In this chapter I have detailed the initial efforts I made to optimise the expression and purification of a number of recombinant PXA domains from various homologues and across several species, as well as a small number of PXC domain constructs. In collaboration with Assoc. Prof. Mike Henne and his laboratory at the University of Texas I have demonstrated a successful method to express and purify the PXA domain from *Oryzias latipes* (Japanese rice fish), leading to some initial crystallisation hits that I hope in the future will lead to a high-resolution structure of this unique protein module.

7.2 Materials and methods

7.2.1 Construct design

The bacterial constructs used in this chapter are from mouse, human, zebrafish and *Chaetomium thermophilum* (thermophilic filamentous fungi) (**Table 10**). Some of these constructs including construct numbers 1, 6-7, 11-12, 15, 17 were cloned by previous lab members. I cloned constructs 2 and 16 (mSNX13 PXC and mSNX25 PXA). All remaining human, *C.thermophilum* and zebrafish constructs (3-5, 8-10, 13-14, 18-19) containing PXA and RGS domains or PXA domain alone, were synthesised and cloned by Genscript (USA) in pGEX-4T-2. The domain boundaries were decided based on the previous studies

published by our lab that demonstrates the SNX-RGS protein domain architecture as well as the secondary structure predictions (38).

Bacterial constructs 20-22 were generated and kindly sent to us by Ryan Feathers from the lab of Assoc. Prof. Mike Henne (UTSW, USA). These constructs are from *O.latipes* and are cloned in pET28(a) with a His-SUMO tag.

Table. 10: Summary of PXA domain constructs used for recombinant bacterial protein expression and purification.

No	Construct	Vector/ Tag	Cleavage site
	SNX13		
1	mSNX13 (97-284) PXA	pGEX-4T-2/ N-term GST	Thrombin
2	mSNX13 (772-957) PXC	pGEX-4T-2/ N-term GST	Thrombin
3	hSNX13 (97-496) PXA-RGS	pGEX-4T-2/ N-term GST	Thrombin
4	ct19887 (116-578) PXA-RGS	pGEX-4T-2/ N-term GST	Thrombin
5	zSNX13 (114-570) PXA-RGS	pGEX-4T-2/ N-term GST	Thrombin
	SNX14		
6	mSNX14 (126-303) PXA	pGEX-4T-2/ N-term GST	Thrombin
7	mSNX14 (768-937) PXC	pGEX-4T-2/ N-term GST	Thrombin
8	hSNX14 (130-304) PXA	pGEX-4T-2/ N-term GST	Thrombin
9	ct18515 (191-373) PXA	pGEX-4T-2/ N-term GST	Thrombin
10	zSNX14 (129-303) PXA	pGEX-4T-2/ N-term GST	Thrombin
	SNX19		
11	mSNX19 (87-271) PXA	pGEX-4T-2/ N-term GST	Thrombin
12	mSNX19 (822-997) PXC	pGEX-4T-2/ N-term GST	Thrombin
13	hSNX19_(95-272) PXA-OPT	pGEX-4T-2/ N-term GST	Thrombin
14	zSNX19_(89-262) PXA-OPT	pGEX-4T-2/ N-term GST	Thrombin
15	mSNX19 (93-274) PXA	Ligation Independent Cloning (LIC)/ N-term HIS6	TEV
	SNX25		
16	mSNX25 (1-164) PXA	pGEX-4T-2/ N-term GST	Thrombin
17	mSNX25 PXC	pGEX-4T-2/ N-term GST	Thrombin
18	hSNX25 (1-401) PXA-RGS	pGEX-4T-2/ N-term GST	Thrombin
19	zSNX25 (121- 527) PXA-RGS	pGEX-4T-2/ N-term GST	Thrombin
	In Collaboration with Assoc. Prof. Mike Henne (UT)		
20	olSNX14 PXA (129-228)	pET28(a)/ N-term His-SUMO	SUMO protease (ULP)
21	olSNX14 PXA (129-264)	pET28(a)/ N-term His-SUMO	SUMO protease (ULP)
22	olSNX14 PXA (129-306)	pET28(a)/ N-term His-SUMO	SUMO protease (ULP)

7.2.2 Transformation and expression of SNX-RGS proteins in *E. coli*

Plasmid DNA coding SNX-RGS domains (except constructs 20-21) were transformed into *E. coli* strain Rosetta cells for expression with ampicillin and chloramphenicol (Amp+/Cm+) antibiotics. A single colony was picked from the LB agar plate and inoculated into 50 mL LB^{2+} with ampicillin (0.1 mg/mL) and chloramphenicol (0.1 mg/mL). It was then incubated at 37°C with shaking overnight. The following day, 30 mL from the overnight culture was used to inoculate 1 L LB media containing ampicillin (0.1 mg/mL) and chloramphenicol (0.1 mg/mL) and incubated at 37°C. Cells were grown to an optical density (OD) of 0.1-0.2 at 600 nm absorbance and induced with 0.2 mM IPTG. This induced bacterial culture was incubated on the shaker overnight at 16°C until the cells reach an O.D of 3.0 (~ 24 h). Cells were harvested by pelleting the culture using Beckman rotor JLA 8.1000 at an RPM of 4000 for 30 min at 4°C. Pellets were resuspended in 10 mL lysis buffer (50 mM Tris pH 8.0, 500 mM NaCL, 5% glycerol, 1 mM DTT, 0.1 mg/ml benzamidin, 0.1 mg/ml DNase) per litre of culture, and stored at –80°C until required or are directly subjected to cell disruption to continue with protein purification.

Constructs 20-22 were sent to us as agar stabs in NiCo21 (DE3) *E.coli* competent cells. The agar stab was inoculated into 50 mL LB^{2+} with kanamycin (0.1 mg/mL). The following day, 50 mL from the overnight culture was used to inoculate 1 L terrific broth (TB) containing kanamycin (0.1 mg/mL) and 70 µM oleic acid incubated at 37°C. Cells were grown to an optical density (OD) of 0.6- 0.7 at 600 nm absorbance and induced with 0.2 mM IPTG. This induced bacterial culture was incubated on the shaker overnight at 16°C until the cells reach an O.D of 3.0 (~ 24 h). Cells were harvested by pelleting the culture using Beckman rotor JLA 8.1000 at an RPM of 4000 for 30 min at 4°C. Pellets were resuspended in 10 mL lysis buffer (50 mM Tris pH 8.0, 500 mM NaCL, 5% glycerol, 1 mM β-octylglucoside, 0.1 mg/ml benzamidin, 0.1 mg/ml DNase) per litre of culture.

7.2.3 Recombinant protein purification of SNX-RGS

The stored resuspended pellets were thawed at room temperature. The thawed pellets were then subjected to cell disruption at 27 kpsi at 5ºC. Samples were collected from lysed cells and dissolved into SDS loading dye and checked by SDS-PAGE analysis. The cells were then centrifuged using JA25.50 rotor at 18,000 RPM for 30 min. The supernatant was incubated with Glutathione-S-Sepharose resin for 30 min before eluting the flow through. After incubation, the flow through was removed and the resin was washed with 5 column volumes of buffer A (50 mM Tris pH 8.0, 500 mM NaCl, 5% glycerol, 1 mM DTT). After washing, beads were mixed with 5 mL of buffer A and 100 U of Thrombin protease

(Sigma-Aldrich) and the mixture was incubated at room temperature overnight to remove the GST tag. The next morning, the cleaved protein was collected in the flow through. SDS-PAGE was used to analyse the samples from each step of the purification. Used resin was regenerated using 40 mM reduced L-glutathione followed by column regeneration using 6 M guanidine HCl with 0.2 M acetic acid, which was then equilibrated using buffer A.

For the constructs 20-22, the purification was done using the buffer 50 mM Tris pH 8.0, 500 mM NaCl, 5% glycerol, 1 mM β-octylglucoside. After cell disruption, the cells were centrifuged and the supernatant was collected. The supernatant was incubated with Talon resin for 30 min before eluting the flow through. After incubation, the flow through was removed and the protein was eluted using 50 mM TRIS pH 8.0, 500 mM NaCl, 5% glycerol, 1 mM β-octylglucoside, 250 mM imidazole. The eluted protein was then mixed with 2 mg of SUMO protease (ULP) and incubated overnight in the cold room, while the mixture was dialysed into 50 mM TRIS pH 8.0, 150 mM NaCl, 5% glycerol, and 1 mM β-octylglucoside. The next morning, the cleaved protein was incubated again with Talon resin to remove uncleaved His-SUMO SNX14 PXA protein. After incubation, the cleaved protein was collected in the flow through. The eluted cleaved protein was then subjected to gel filtration using a Superdex 75 column (10/300) which was equilibrated in 50 mM Tris pH 8.0, 150 mM NaCl, 5% glycerol, and 1 mM β-octylglucoside. SDS-PAGE was used to analyse the samples from each step of the purification. The successfully eluted and pooled protein was then dialysed into 20 mM Tris pH 8.0, 150 mM NaCl for further crystallisation experiments.

7.2.4 CD spectroscopy

The structure of proteins 20-22 in solution (0.2 mg/ml) was determined in a quartz cuvette of 1.0 mm path length using a standardized methodology for CD spectropolarimeter (Jasco J-810). The blank was 10 mM sodium phosphate pH 7.4. Each CD spectrum consisting of the ellipticity and absorbance values, was obtained over a wavelength range from 190 to 250 nm, at a scan rate of 50 nm/min and a response time of 0.25 sec. Each spectrum represented an accumulation of 3 scans. The data was analysed using the spectra analysis program in the spectra manager software.

7.2.5 Construct design for The BAC-TO-BAC® Baculovirus Expression System

The BAC-TO-BAC® Baculovirus Expression System (Invitrogen) was used to generate recombinant baculovirus stocks for the expression of hSNX13 PXA-RGS, ct19887

(SNX13) PXA-RGS, hSNX14 PXA, ct19557 (SNX14) PXA and hSNX25 PXA. An overview of the steps involved in this procedure is shown schematically in **Fig. 64**.

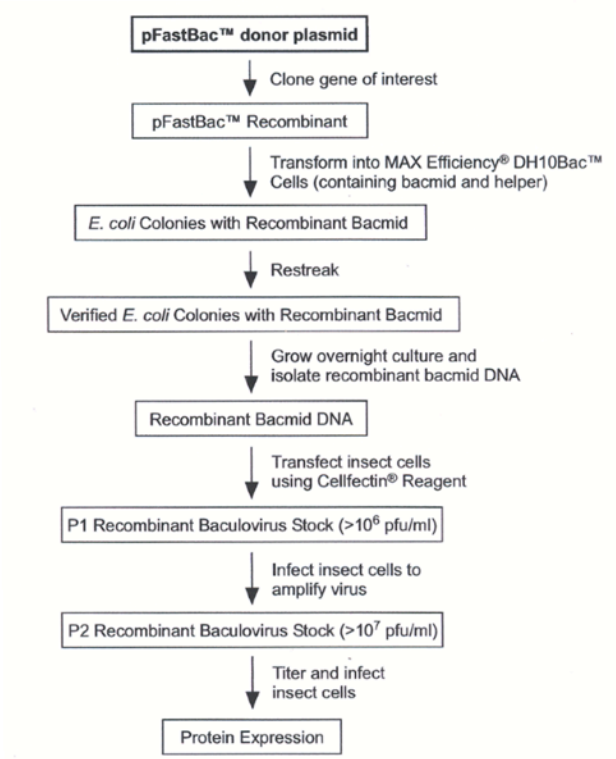


Figure 64. *Overview of the steps involved in the generation of recombinant baculoviruses using the BAC-TO-BAC® Expression System.*

7.2.6 Design of oligonucleotides

To facilitate purification, human and *C.thermophilum* SNX13 and SNX14 constructs were expressed as hexahistidine (6His)-tagged fusion proteins. Primers were designed to generate PCR products that could be cloned directly into the BamHI and XhoI sites of the multiple cloning site of the pFastBac1 vector. For the generation of N-terminal His-tagged proteins, the 5' primers have the following features: i) three extra 5' bases for stability as recommended by New England Biolabs (NEB); ii) a BamHI restriction site; iii) an additional methionine codon to initiate translation, followed by a serine residue to promote stability; and iv) six consecutive histidine residues. The 3' primers have two stop codons to ensure translation termination followed by a XhoI restriction site. C-terminal His-tagged proteins were generated as follows: The 5'-primers were designed to have a BamHI restriction site, followed by a start codon (ATG) and 3'-primers included six consecutive histidine residues, followed by two stop codons and a XhoI restriction site.

7.2.7 Amplification of DNA by PCR

DNA fragments of the required length were generated by the polymerase chain reaction (PCR) Standard PCR reactions contained 10 μl 5X Phusion HF buffer, 500 ng Template

DNA, 1 μl dNTP mixture (10 mM), 10 μM each of forward and reverse primers and 0.5 μl of Phusion DNA polymerase (NEB). The reaction volume was adjusted to 50 μl with sterile MilliQ™ H_2O. Thermocycling conditions used to produce SNX13 and SNX14 gene products were as follows:

Step	Time	Temperature	Cycles
Initial Denaturation	30 s	98°C	1X
Denaturation	20 s	98°C	25-35X
Annealing	30 s	60°C	
Extension	2 min	72°C	
Final Extension	7 min	72°C	1X

7.2.8 Purification of PCR fragments

Amplified PCR reactions were purified using a Nucleospin Gel and PCR Clean-up Kit (Machery Nagel) following the manufacturer's protocol. PCR products were eluted in 35 μls of EB buffer.

7.2.9 Analysis of DNA by agarose gel electrophoresis

DNA was characterised by gel electrophoresis to confirm purity and molecular weight. Agarose gels of 1% (Molecular Biology Grade, Progen) were electrophoresed at 100 V for 30 minutes in 1X TAE buffer. DNA was loaded in either DNA Sample Buffer or 50% glycerol. Molecular weight markers used were the GeneRuler™1 kB and 100 bp DNA ladders (Bioline and NEB). DNA was visualised by SYBR Safe staining.

7.2.10 Restriction-enzyme digestion of PCR products and pFastBac1 DNA

Restriction enzymes and digestion buffers were used according to the manufacturer's instructions (NEB). Vector (pFastBac1) and PCR products were double digested with BamHI and XhoI in Cutsmart buffer. Reactions were incubated at 37°C for 2 h. After digestion was complete, vector DNA was dephosphorylated to prevent self-ligation. Antarctic alkaline phosphatase was added directly to the digestion mixture that was incubated for a further 30 min at 37°C.

7.2.11 Ligation of DNA

The calculated molar ratio of vector to insert DNA in ligations was generally 1:3. Ligation reactions were prepared in a volume of 20 μL containing 50-100 ng of vector DNA, a corresponding amount of insert, 5X Ligation Buffer and 1 unit of bacteriophage T4 DNA ligase (NEB). Reactions were incubated for 4 h at 16°C. An aliquot (2-10 μL) of the reaction mixture was used to transform into competent *E. coli* DH5α cells.

7.2.12 Transformation of plasmid DNA and ligation mixtures into *Escherichia coli*

Plasmid DNA or ligation products were transformed into chemically competent DH5α cells. Competent cells (100 μl aliquots) were thawed on ice. Plasmid DNA (1-10 ng) or ligation mixtures (2-10 μL) were gently added and mixture transferred to a chilled 14 ml polypropylene culture tube (Falcon) and incubated on ice for 30 min. The cell/DNA mix was heat-shocked by incubation in a heating block at 42°C for 60 s and immediately cooled on ice for 2 min. A pre-warmed aliquot of LB medium (900 μl) was added to each culture. Samples were incubated with shaking (250 rpm) at 37°C for 1 h. Aliquots (50-200 μl) from each transformation culture were plated onto LB agar plates containing 100 μg/ml ampicillin (or other appropriate antibiotics) and incubated overnight at 37 °C.

7.2.13 Preparation and purification of plasmid DNA

A single colony containing plasmid DNA of interest was used to inoculate 3 ml of 1X LB broth (containing 100 μg/ml ampicillin). Cultures were incubated with vigorous shaking at 37°C for 12-16 h. Cells were collected by centrifugation (JS 5.3, 5000 rpm, 10 min, RT) and the pellet was processed according to the Nucleospin Plasmid kit manufacturers protocol (Machery Nagel). Purified DNA was digested as described above to confirm the presence of cloning junctions and an insert of correct size before sequencing.

7.2.14 Nucleotide sequence analysis

Purified products were sent the Australian Genome Research Facility (Brisbane) for automated analysis. DNA sequences were analysed using MacVector (version 12.0.6) and ExPASy.

7.2.15 Transposition

Transposition Media plates [LB plates containing containing kanamycin (50 μg/ml), gentamicin (7 μg/ml), and tetracycline (10 μg/ml), 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal) (200 μg/ml) and isopropyl-β-D-thiogalactopyranoside (IPTG) (40 μg/ml)] were prepared. DH10Bac™ MAX EFFICIENCY competent cells (Life

Technologies, Invitrogen) were thawed on ice and 50 µl aliquots were dispensed into chilled 14 ml polypropylene culture tubes (Falcon 2059). Recombinant donor plasmid (approximately 1 ng diluted in 5 µl sterile H_2O) was gently added to the cells. The mixture was incubated on ice for 30 min, transferred to a 42°C heating block, heat-shocked for 45 s and immediately chilled on ice for 2 min. SOC medium (900 µl) was added and the culture was incubated at 37°C with medium agitation (200 rpm) for 4 h. Aliquots (50, 100 and 200 µl) were plated onto Transposition Media plates and incubated for 48 h at 37°C. A positive control (pUC19 DNA) was used to assess the transformation efficiency of the DH10Bac™ cells. White *E. coli* colonies containing recombinant bacmid were generally larger and readily distinguishable from non-recombinant blue colonies after incubation at 37°C for 24 h. Before isolating recombinant bacmid DNA, candidate colonies were re-streaked onto fresh Transposition Media plates to verify their phenotype.

7.2.16 Purification of recombinant bacmid DNA

A single isolated bacterial colony, confirmed as having a white phenotype on plates containing Bluo-gal and IPTG, was inoculated into LB medium (3 mL) supplemented with 50 µg/mL kanamycin, 7 µg/ml gentamicin and 10 µg/ml tetracycline. Cultures were grown at 37°C with orbital shaking (250-300 rpm) to stationary phase (up to 24 h). Bacterial cells were recovered by centrifugation (5000 rpm, 5 min, RT) and the pellet was re-suspended in 300 µl Resuspension Buffer (solutions from Nucleospin Plasmid kit - Machery Nagel). Lysis Solution (300 µl) was added and the suspension was gently mixed and incubated at room temperature for 5 min, changing from turbid to almost translucent in appearance. Neutralisation Solution (300 µl) was slowly added with gentle mixing and a thick white precipitate of protein and *E. coli* genomic DNA formed. The sample was incubated on ice for 5 to 10 min before centrifugation (14000 rpm, 10 min, 4°C). The supernatant, containing recombinant bacmid DNA, was carefully transferred (to avoid carry-over of any white precipitated material) to a sterile Eppendorf tube containing 800 µl isopropanol. The solution was mixed by gentle inversion, placed on ice for 10 min to precipitate the DNA and centrifuged (14000 rpm, 15 min, RT). The supernatant was removed and 70% ethanol (500 µl) was added to each tube to wash the pellet. The mixture was centrifuged (14000 rpm, 5 min, RT) and washing with 70% ethanol was repeated. The supernatant was removed, ensuring the pellet was not dislodged. The pellet was air-dried at RT for 5 to 10 min, and the DNA was then dissolved in EB buffer (40 µl). Bacmid DNA was stored at –20°C in small aliquots to avoid repeated freeze/thaw cycles that can cause a significant reduction in transfection efficiency.

7.2.17 PCR analysis of recombinant bacmid DNA

As the recombinant bacmid DNA is greater than 135 kb in size, verification of insertion of the gene of interest is difficult using classical restriction endonuclease digestion analysis. It is therefore more effective to use PCR to confirm that the gene of interest has transposed to the bacmid. The pUC/M13 amplification primers are directed at sequences on either side of the mini-attTn7 site within the *lac*Zα-complementation region of the bacmid (**Fig. 65**).

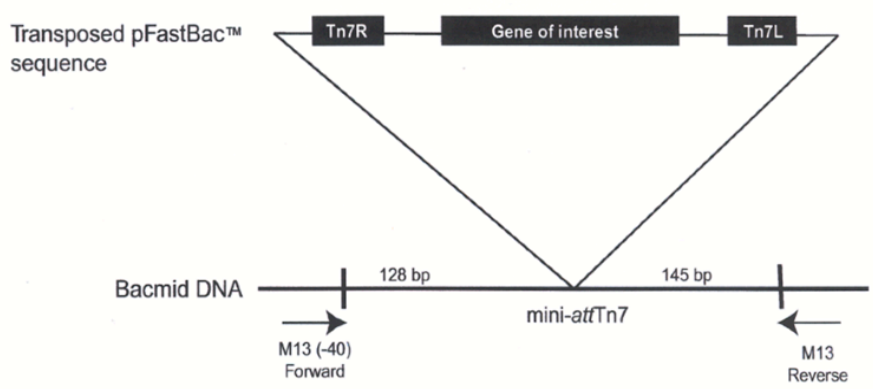


***Figure 65**. M13 forward and reverse priming sites within the pFastBac1 vector. Location of M13 Forward (-40) and M13 Reverse priming sites flanking the mini-attTn7 site within the lacZα-complementation region to facilitate PCR analysis. (Source: https://tools.thermofisher.com/content/sfs/manuals/bactobac_man.pdf).*

These two primers have the following sequences:

(i) pUC/M13 Forward: 5'-CCCAGTCACGACGTTGTAAAACG-3'

(ii) pUC/M13 Reverse 5'-AGCGGATAACAATTTCACACAGG-3'

A PCR reaction was set up for each recombinant bacmid DNA sample containing 10 μl 5X Phusion HF buffer, ~100 ng bacmid DNA, 1 μl dNTP mixture (10 mM), 10 μM each of forward and reverse primers and 0.5 μl of Phusion DNA polymerase (NEB). The reaction volume was adjusted to 50 μl with sterile H_2O. The following parameters were were used for amplification:

Step	Time	Temperature	Cycles
Initial Denaturation	3 min	94°C	1X
Denaturation	45 s	94°C	25-35X
Annealing	45 s	55°C	
Extension	5 min	72°C	
Final Extension	7 min	72°C	1X

An aliquot from each reaction (5-10 μl) was analysed by gel electrophoresis. If transposition has occurred (bacmid transposed with pFastBac1), a PCR product of ~2430 bp plus the size of the SNX13/14 inserts should be generated.

7.2.18 Protein Production and Insect cell culture

Spodoptera frugiperda (Sf9) and *Trichoplusia ni* (BTI-TN-5BI4) (High Five™) insect cells were purchased from Invitrogen. Both cell lines were maintained in suspension culture in shaker flasks, incubated at 27°C and 120 rpm. Sf9 cells were cultured in Sf-900 II SFM (Invitrogen) and High Five™ cells were maintained in Express Five® SFM (Invitrogen) supplemented with L-Glutamine. Cells were routinely cultured until they reached 4 X 10^6 cells/ml and were then diluted to 1 X 10^6 cells/ml to ensure that cells were maintained in exponential growth phase with maximum viability. Sf9 cells were used for all transfections, amplification of viral stocks, plaque assays and protein production. As High Five™ cells generally transfect less efficiently, they were used solely for protein expression.

7.2.19 Transfection of Sf9 cells with recombinant bacmid DNA

Each transfection was performed in duplicate using Sf9 cells in mid-exponential growth phase with a viability of >97%. Approximately 9x10^5 cells were seeded in each 35mm well (of a Nunc 6-well plate) in Sf-900 II SFM (2 ml). Cells were allowed to attach for at least 1 h at 27°C. During this time, two solutions were prepared: Solution A: For each transfection, ~10 µl of purified bacmid DNA was diluted into 100 µl Sf-900II SFM without antibiotics. Solution B: For each transfection ~10 µl Cellfectin® Reagent was diluted into 100 µl Sf-900II SFM without antibiotics. Cellfectin® Reagent is a 1:1.5 (M/M) liposome formulation of the cationic lipid N,NI,NII,NIII-Tetramethyl-N,NI,NII,NIII-tetrapalmitylspermine (TM-TPS) and dioleoyl phosphatidylethanolamine (DOPE) in membrane-filtered water. These solutions were combined, mixed gently and incubated for 45 min at RT. Sf-900 II SFM (800 μl) was added to each tube containing the lipid-DNA complexes and diluted lipid-DNA complexes were gently overlaid onto the cells. Cultures were incubated for 5 h at 27°C. The transfection mixtures were gently removed by aspiration and 3 ml of Sf-900 II SFM (containing antibiotics/antimycotic) was added. Cells were incubated at 27°C for a further 72 h. Transfection supernatants containing recombinant baculovirus were harvested by centrifugation (3000 rpm, 5 min, RT) and duplicate transfection supernatants were pooled. This initial transfection viral stock (designated P1) was used as the inoculum to amplify recombinant baculovirus in Sf9 cells.

7.2.20 Amplification of initial transfection (P1) baculoviral stock to generate P2 intermediate and P3 high titre stocks

The P1 viral stock generated by each transfection was used to infect a monolayer culture to generate an intermediate P2 viral stock. Sf9 cells were grown in suspension culture to mid-exponential phase with >97% viability. Cells were seeded in T-175 tissue culture flasks (Nunc) at a density of 3 x 10^7 cells/flask in 15 ml Sf-900 II SFM and allowed to adhere for 1 h. Attachment and confluency was assessed by examination of the monolayer using an inverted microscope. It was assumed that the titre of the initial viral stock obtained from transfecting Sf9 cells generally ranges from 1 x 10^6 to 1 x10^7 plaque forming units pfu/ml (as outlined the Invitrogen Guide to Baculovirus Expression Vector Systems (BEVS) and Insect Cell Culture Techniques) and it is recommended that infection using a multiplicity of infection (MOI) of 0.01 to 0.1 will result in approximately 100-fold amplification of the virus. MOI is defined as the number of virus particles per cell. The following formula was used to calculate the volume of viral stock to add to obtain a specific MOI:

$$Inoculum\ required\ (ml) = \frac{\left[MOI\ \left(\frac{pfu}{cell}\right) x\ number\ of\ cells\right]}{titre\ of\ viral\ stock\ \left(\frac{pfu}{ml}\right)}$$

After cells had firmly adhered, media was aspirated from the cell monolayer and viral inoculum was added (1-2 ml of P1 stock diluted into approximately 15 ml Sf-900 II ensuring that the cell monolayer was covered). The virus was allowed to adsorb at RT for 1 h and a further 40-50 ml of Sf-900 II media was added. Cultures were incubated at 27°C for 48-72 h. Optimal harvest times can vary and were determined for each baculoviral construct. The supernatant from this amplification was harvested, clarified in a benchtop centrifuge (3000 rpm, 5 min, RT) and 5 ml of this amplified supernatant was added directly to Sf9 cells in suspension culture (400 ml) at a density of 2 x 10^6 cells/ml. This culture (designated P3), was incubated at 27°C and 120 rpm and recombinant baculovirus was harvested 3 days post-infection by centrifugation (3000 rpm, 5 min, RT). FBS was added as a preservative to a final concentration of 2% to the clarified viral stocks which were stored at 4°C, protected from light. A small aliquot was also stored at –70°C. This amplification procedure allows production of a high titre P3 viral stock with a titre ranging from 1 x 10^7 to 1 x 10^8 pfu/ml, recommended for expression studies.

7.2.21 Test expressions

The expression of each recombinant protein of interest was initially analysed in static Sf9 cultures. Sf9 cells were grown in suspension culture to mid-exponential phase with >97%

viability. Cells were then seeded in T-175 tissue culture flasks (Nunc) at a density of 3 x 10^7 cells/flask in 15 ml Sf-900 II SFM and allowed to adhere for 1 h. Approximately 2-3 ml of the intermediate P2 viral stock for each recombinant protein was used to infect each the monolayer culture. The virus was allowed to adsorb at RT for 1 h and a further 40-50 ml of Sf-900 II media was added. Cultures were incubated at 27°C for 72 h.

7.2.22 Harvesting and analysis of recombinant expression

Recombinant SNX13 and SNX14 proteins accumulated intracellularly. Infected insect cells were harvested (from T-175 flasks) at 72 h post-infection and cells collected by low speed centrifugation (3000 rpm, 10 min, RT). Insect cell pellets were washed once with phosphate-buffered saline (PBS) and then resuspended in 3 ml PBS. Cells were lysed by sonication (typically 1 x 30s burst at 60W using a Misonix Sonicator 3000). The sonication was repeated until microscopic observation of the culture revealed few intact cells. The lysate was clarified by centrifugation (14 000 rpm, 10 min, 4°C). The supernatant and pellets were then analysed by SDS-PAGE, followed by Western blotting. Soluble cytoplasmic proteins are found in the supernatant, with insoluble material retained in the pellet.

7.2.23 Western blotting

Proteins were separated by SDS-PAGE as previously described membrane using the iBlot Dry Blotting System (Life Technologies, Invitrogen). A pre-stained molecular weight marker (BLUltra, GenedireX) was used to monitor the transfer of proteins to the membrane. Membranes were blocked with 5% skim milk powder in PBS for 2 h at RT and then incubated overnight with the primary monoclonal antibody [Penta•His (Qiagen)] diluted 1:1 000 in PBS-skim milk. After 2 x 5 min washes in PBS, membranes were then incubated for 1 h with a goat anti-mouse-horseradish peroxidase (HRPO) conjugated secondary antibody (Life Technologies, Invitrogen) diluted 1:2000 in PBS-skim milk. Membranes were then washed for 3 x 5 min in PBS to remove unbound conjugate, and developed using a Novex ECL Chemiluminescent Substrate Kit (Life Technologies, Invitrogen). Proteins were detected on X-ray film (Kodak).

7.2.24 Infection of insect cells with recombinant baculoviruses

One large scale infection of Sf9 cells (6 L) with hSNX13 PXA-RGS was performed to generate protein for an investigation of whether SNX13 and SNX14 proteins expressed in insect cells could be purified without the detergent required for proteins produced in the bacterial system. Cells were infected whilst in the mid-logarithmic phase of growth at a

density of 2 x 10^6 cells/ml at a MOI of ~5. Cultures were incubated at 27°C with shaking at 120 rpm. Cells were harvested at 72 h post-infection by centrifugation (3000 rpm, 10 min, RT). Similar purification procedures were then used for both bacterial and insect cells. As infection conditions in insect cells can vary, optimisation experiments will need to be performed to determine optimal infection parameters, particularly cell line, MOI and establishment of the time course of protein production for each recombinant baculovirus.

7.3 Results

7.3.1 SNX-RGS bacterial protein test expression and purification trials

When I started working on the PXA constructs from SNX-RGS proteins, the only known information was this domain could be a protein-protein interaction region. These purification trials were done based on the available information, treating the proteins as typical small globular domains that should be soluble in standard ionic buffer solutions.

Test expressions of all constructs using 1 L cultures were performed for the constructs 1-2, 6-7, 11-12, 15, 16-17 (**Table 10**). Figures show that GST-tagged mSNX13 PXA, mSNX13 PXC (**Fig. 66A**), mSNX14 PXA, mSNX14 PXC (**Fig. 66B**), mSNX25 PXA, mSNX25 PXC (**Fig. 66C**), mSXN19 PXA and mSNX19 PXC (**Fig. 66D**) all exhibit a significant degree of soluble expression (protein remaining in the soluble supernatant) and bind to the glutathione-resin. Even though it appears the cleaving of GST tag is successful, all of the tested proteins failed to elute off the GST column after the thrombin cleavage, suggesting poor folding or solubility for the isolated PXA and PXC domains.

Expression of each of these constructs suggested that significant amounts of soluble GST-fusion proteins could be produced; however the lack of elution from glutathione affinity matrix was puzzling. This suggested possible aggregation or mis-folding of these constructs.

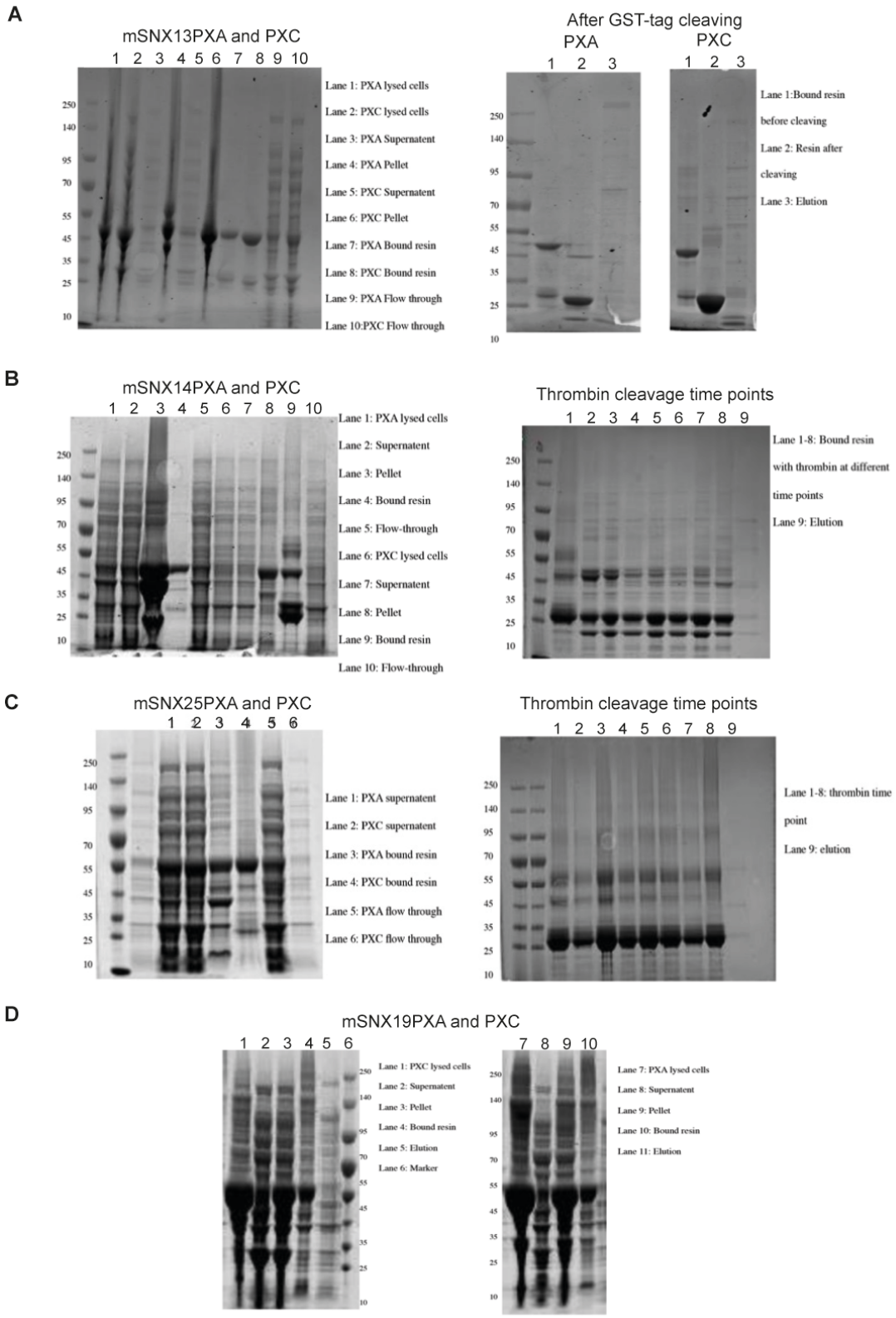


Figure 66. *SDS-PAGE analysis of affinity chromatography purification of GST-tagged* ***(A)*** *mSNX13 PXA and PXC (MW=46 kDa),* ***(B)*** *mSNX14 PXA and PXC (MW=47 kDa),* ***(C)*** *mSNX25 PXA and PXC (MW=56 kDa), and* ***(D)*** *mSNX19 PXA and PXC (MW=52 kDa) showing the presence of expressed protein at their respective expected molecular weights. Even though the fractions of GST resin showed their presence, these proteins failed to elute from glutathione sepharose after the thrombin cleavage of the GST tags.*

We next decided to test additional constructs of this protein from different species (constructs 3-5, 8-10, 13-14, 18-19). The other species were human, zebrafish and *Chaetomium thermophillum*. Some of these constructs had both PXA and RGS domains to test whether the presence of an additional domain might provide extra stability or solubility. Even though these constructs were all also highly expressed, as for the mouse constructs the elutions were again unsuccessful. Examples of two constructs are shown in **Fig. 67A**.

Since all the proteins were expressed but failed to elute after thrombin cleavage to remove GST (or using glutathione elution of the GST-tagged proteins (not shown), I speculated the problem might be protein misfolding or hydrophobicity leading to aggregation, making them difficult to elute it off the column in typical aqueous buffer. Different buffer conditions were therefore tried to see if these could help reduce any non-specific binding of the proteins to the glutathione sepharose. When 0.2% Triton X-100 was added to the elution buffer the GST-tagged fusion proteins began eluting off the column, suggesting the non-specific binding of a hydrophobic protein to the GST resin. As seen in the **Fig. 67B,** the elution with the triton detergent helped regardless of the buffer condition.

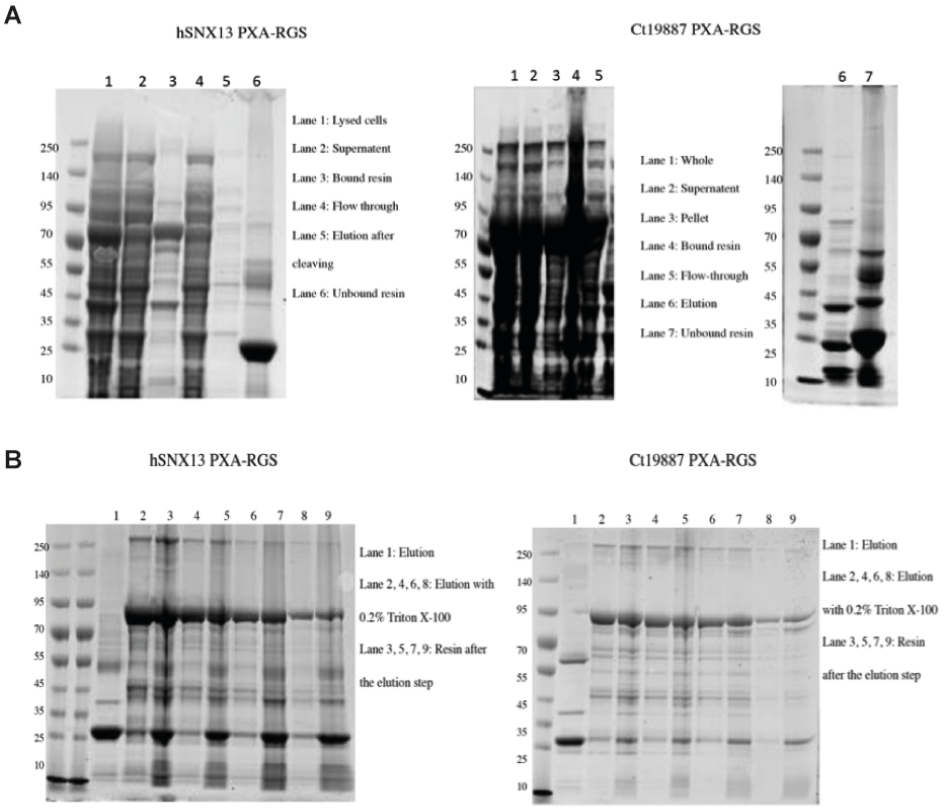


Figure 67. *SDS-PAGE gel analysis of fractions from* ***(A)*** *affinity chromatography of GST-tagged hSNX13 PXA-RGS (MW=82 kDa) and ct19887 PXA-RGS (MW=80 kDa) showing the presence of expressed protein.* ***(B)*** *Elution of hSNX13 PXA-RGS and ct19887 PXA-RGS using 0.2% Triton X-100 with buffer conditions containing different salt concentrations (50 mM Tris pH 8.0, 30 mM L-glutathione reduced, with 500 mM NaCl, 200 mM NaCl, 150 mM NaCl, and 100 mM NaCl).*

I next attempted to cleave the GST tag using thrombin in-solution rather than on the beads, which did yield the protein of interest, albeit in a relatively impure state (**Fig. 68, right**). The next hurdle was to remove the Triton using biobeads from the eluted protein, which was unfortunately not successful in yielding a pure and homogeneous hSNX13 PXA-RGS protein. The other protocol I tried to improve the solubility was by co-expressing with molecular chaperons (169), which also proved to be unsuccessful in our case. Co-expression as well as co-purification with a potential binding partner, CARD domain from ARC (50) was also performed with no success.

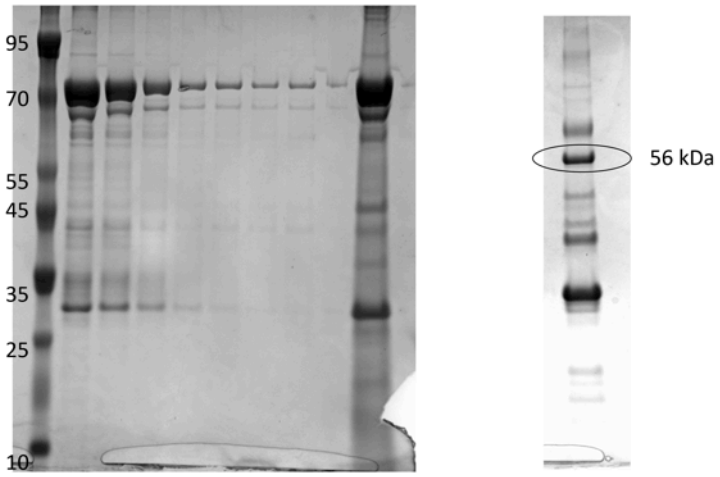


***Figure 68**. **(Left)** SDS-PAGE showing the elution of hSNX13 PXA-RGS protein (around 82 kDa) using 0.2% Triton X-100; **(Right)** SDS-PAGE showing the presence of the cleaved hSNX13 PXA-RGS protein of interest after cleaving off the GST-tag. GST is present at slightly higher than 25 kDa.*

While it was speculative at this stage, these results from the recombinant expression of the PXA and PXC domains of human, mouse and *C. thermophilum* SNX-RGS proteins began to suggest that these domains might be lacking a key co-factor or ligand required for their stability. When Mike Henne and his group published the potential role of SNX-RGS proteins in ER-endosome interactions (59), I began to speculate that these unknown PXA and PXC domains could regulate lipid interactions or lipid transfer analogous to other proteins such as the ORP family members (170). Thus my results suggested to me that these PXA proteins are highly hydrophobic and may require a ligand such as a lipid or another molecule that is just not present in bacteria. If so it is possible that bacterial protein expression is unable to provide a potential lipid co-factor required for protein stability. Hence the next step included further exploration of expression and purification of these SNX-RGS constructs in insect cells with the hope that this will allow

the proteins to be correctly folded and to incorporate any potential native ligands not present in bacteria.

Expression of SNX-RGS constructs in insect cells has been performed with assistance from Dr. Suzanne Norwood. I performed all of the cloning experiments and transpositions, and Dr. Suzanne Norwood performed transfections, test expression and assisted with Western-blotting.

Most PCR reactions generated one dominant fragment of the correct size and the amplified inserts were digested and directly ligated into the pFastBac1 expression vector without requiring further purification (**Fig. 69**).

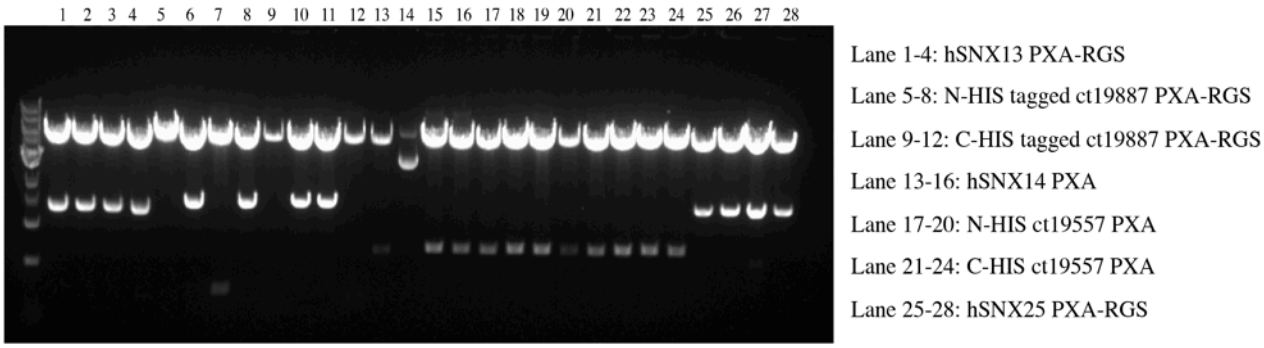


Figure 69. *Restriction digestion on ligation colonies showing the presence of inserts.*

After the transposition reaction had been successfully performed, high molecular weight recombinant bacmid DNA was isolated, purified and used to transfect insect cells to generate recombinant baculoviruses, which were used for preliminary expression experiments. The initial transfection baculoviral stock (P1) was only amplified twice (to avoid the introduction of mutations) and this generated intermediate (P2) and high titre (P3) stocks that were used to infect insect cells for initial test expressions and large-scale expression. In test expressions, the recombinant hSNX13 PXA-RGS construct expressed well compared to the ASIC channel construct used as a control, but the *C. thermophilum* SNX13 homologue was not (**Fig. 70**). Both PXA and PXC domains from *C. thermophilum* SNX14 proteins were also expressed and soluble. Despite these promising results I was not able to purifiy significant amounts of soluble homogeneous protein following large-scale expressions. However this was only attempted once and bears repeating in the future.

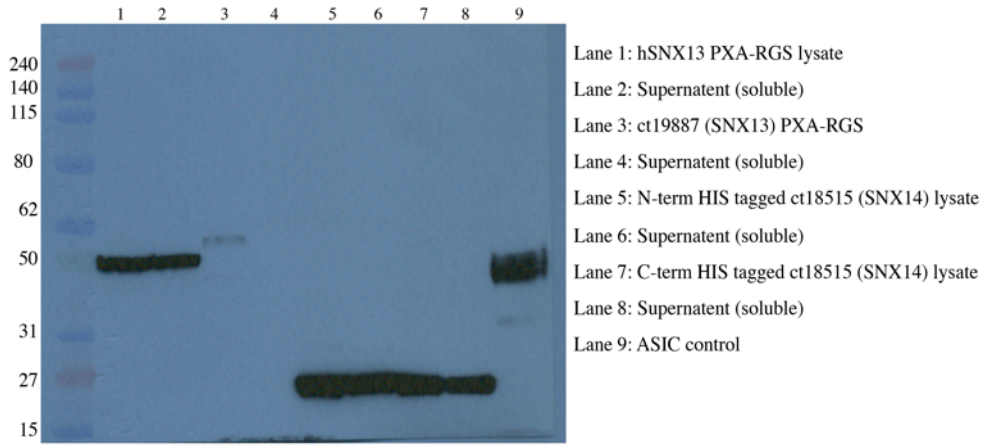


Figure 70. *Western blot of the test expressions demonstrating the presence of soluble proteins.*

7.3.2 Recombinant protein expression and purification of constructs 20-22

The constructs 20-22 from *O. latipes* SNX14 (will be addressed as OL1, OL2 and OL3 herein) were cloned by Ryan Feathers and a gift from the lab of Assoc. Prof. Mike Henne (UTSW). These three bacterial constructs were designed based on *in silico* predictions of different secondary structure elements in the domain. OL1 is missing 77 AAs, whereas OL2 is missing 37 AAs from the C-terminal α-helical region when compared to the full-length construct OL3 (**Fig. 71**).

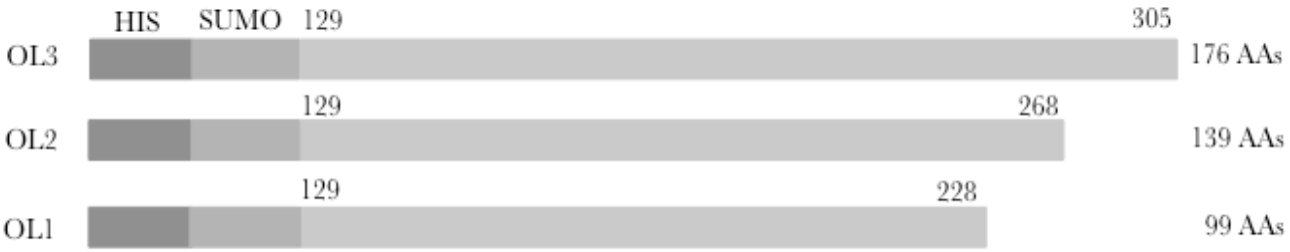


Figure 71. *Three SNX14 PXA (O. latipes) constructs (OL1, OL2 and OL3) designed based on the predicted sequence-based protein solubility.*

Purification of constructs OL1, OL2 and OL3 was performed using affinity chromatography with Talon resin followed by gel filtration using Superdex 75 (10/300) column. Samples from every step of purification were collected for SDS-PAGE and were analysed for the presence of successfully purified proteins of interest (**Fig. 72**). Fractions that are found to be relatively pure were pooled together for further experiments. After the gel filtration, the yield of construct 1 was around 20 mg/ml. The constructs 2 and 3 failed to successfully express soluble proteins.

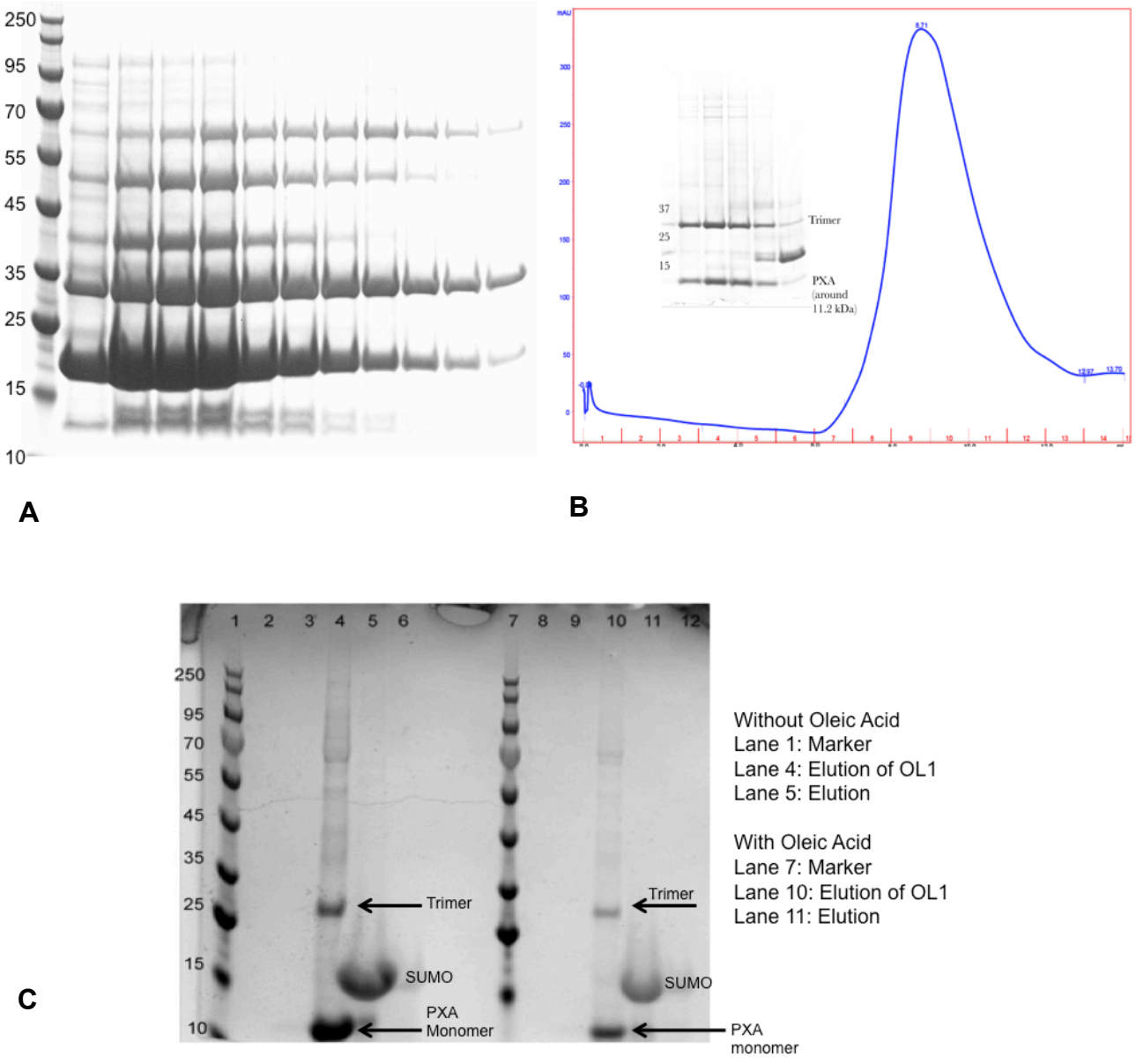
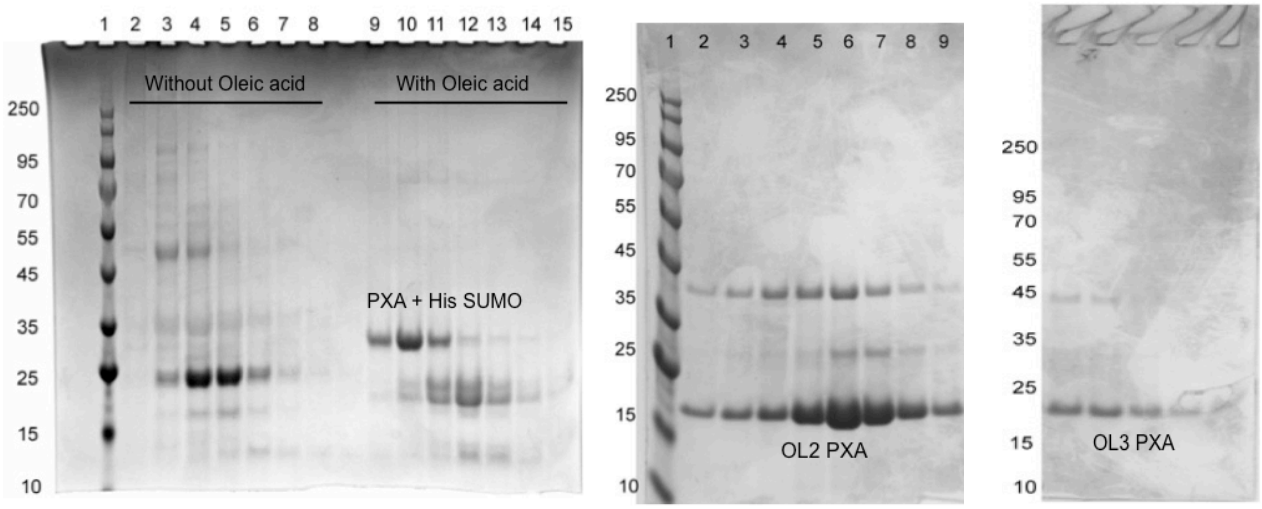


***Figure 72**. **(A)** SDS-PAGE analysis of elution fractions from the Affinity chromatography of His-SUMO (12 kDa) tagged construct OL1 SNX14 PXA protein demonstrating the presence of expressed soluble protein at the expected molecular weight around 23 kDa; **(B)** FPLC chromatogram of cleaved SNX14 PXA construct OL1 (11 kDa) and SDS-PAGE gel of fractions corresponding to the major peak; **(C)** SDS-PAGE gel showing FPLC purified tagless construct OL1 that was expressed without and with oleic acid.*

I also decided to test the addition of Oleic acid (70 μM) to the culture medium during the protein expression since the PXA domain was found to bind to long chain fatty acids (Mike Henne, unpublished). Oleic acid helped protein expression and constructs OL2 and OL3 expressed soluble proteins and seemed to be less aggregated, even though the

yields of construct OL3 was low (**Fig. 72C and 73**). Using analytical ultracentrifugation and crosslinking, the laboratory of Mike Henne has confirmed the presence of multimeric species in the soluble OlSNX14 protein.



A

Figure 73*. (A) SDS-PAGE analysis of elution fractions from the affinity purification of His-SUMO (12 kDa) tagged SNX14 PXA (around 28 kDa) construct OL2 expressed without and with oleic acid; The expression of OL2 without oleic acid is very poor, whereas the gel demonstrate the presence of soluble protein when oleic acid is introduced during protein expression (B) SDS-PAGE gel of the FPLC purified tagless construct OL2 (around 16 kDa) (expressed with OA); (C) SDS-PAGE gel of the FPLC purified tagless OL3 (around 20.3 kDa) (expressed with OA).*

Construct OL2 yielded around 12 mg/ml whereas construct 3 yielded only 7 mg/ml after concentrating down the eluted and pooled fractions. Thus when expressed with low concentrations of oleic acid and purified with low concentrations of β-octylglucoside detergent, all three constructs of SNX14 PXA (*O.latipes*) were found to be soluble. These proteins were also folded and showed the presence of α-helical structure as expected from secondary structure predictions when tested using CD spectroscopy (**Fig. 74**).

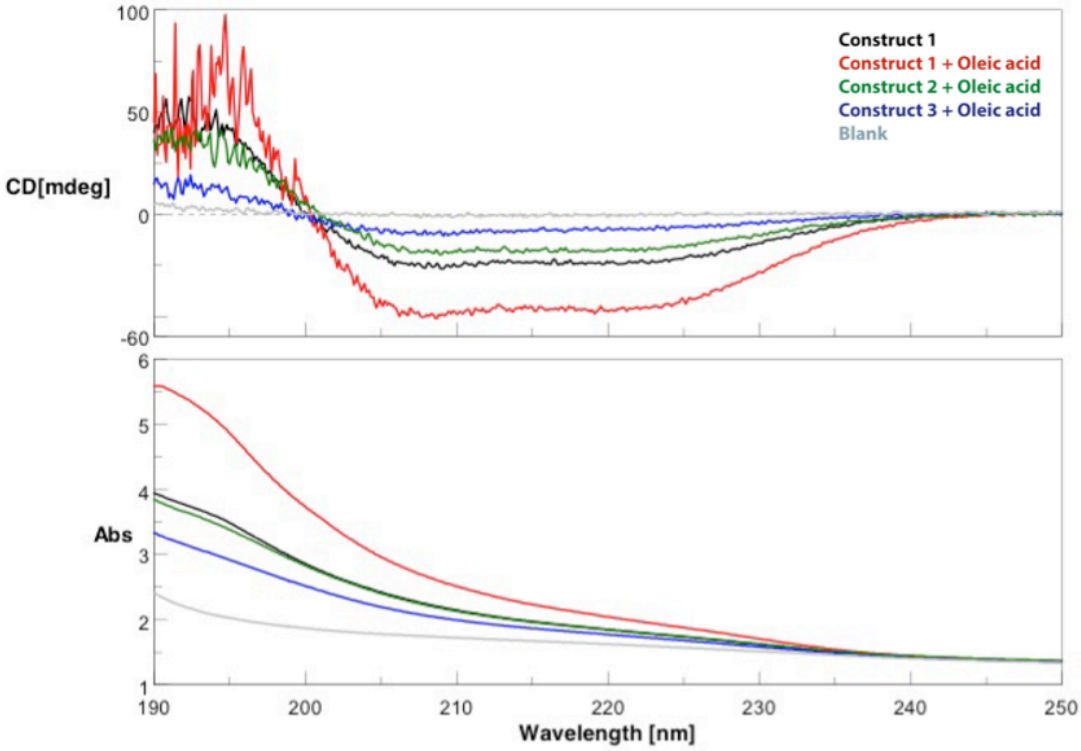


Figure 74. *CD spectra of OL1, OL2 and OL3 constructs of SNX14 PXA (O.latipes) showing the α-helical structure as expected thus confirming the soluble proteins are properly folded.*

7.3.3 Crystallisation trials of SNX14 PXA (*Oryzias latipes*)

All three constructs of SNX14 PXA from *O. latipes* produced needle crystals in both 96-well crystallization screens and also 24-well optimization grids. The conditions that gave these crystals are listed in **Table 11**. Low concentrations of oleic acid were also added to the purified PXA protein to observe any improvements in the crystals. Some of the small needle crystals obtained is shown in **Fig. 75**. Clearly these crystals are currently too small for high-resolution structure determination, but are promising leads for further optimization of both sample preparation and crystallization conditions.

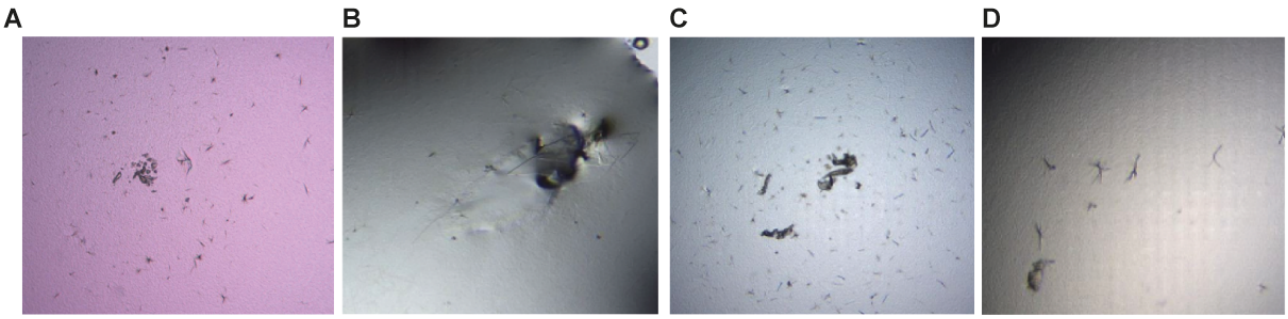


Figure 75. *Crystals of from 24-well optimization plates (**A**) Needle crystals of OL1 SNX14 PXA; (**B-C**) Crystals of OL2 SNX14PXA; (**D**) Crystals of OL3 SNX14PXA*

Table 11. Some of the crystallisation reservoir conditions that produced positive crystal hits of OL1, OL2 and OL3 SNX14 PXA (*O. latipes*)

1. 10% (v/v) 2-propanol, 0.1 M HEPES pH 7.5, 0.2 M NaCl
2. 10% (w/v) PEG-3000, 0.1 M CHES pH 9.5
3. 20% (w/v) PEG-1000, 0.1 M Na/K phosphate pH 6.2, 0.2 M NaCl
4. 10% (w/v) PEG-3000, 0.1 M imidazole pH 8.0, 0.2 M Li_2SO_4
5. 1.26 M $(NH4)_2SO_4$, 0.1 M HEPES pH 7.5
6. 10% (v/v) 2-propanol, 0.1 M cacodylate pH 6.5, 0.2 M $Zn(OAc)_2$
7. 50% (v/v) ethylene glycol, 0.1 M imidazole pH 8.0
8. 50% (v/v) ethylene glycol, 0.1 M HEPES pH 7.5, 0.2 M Li_2SO_4
9. 40% (v/v) ethylene glycol, 0.1 M Tris pH 7.0
10. 30% (v/v) PEG-200, 0.1 M CAPS pH 10.5, 0.2 M $(NH_4)_2SO_4$
11. 10% (v/v) 2-propanol, 0.1 M HEPES pH 7.5, 0.2 M NaCl
12. 30% (w/v) PEG-3000, 0.1 MCHES pH 9.5
13. 20% (w/v) PEG-8000, 0.1 M CAPS pH 10.5, 0.2 M NaCl

7.4 Discussion

Deciphering the molecular mechanisms by which the SNX-RGS proteins interact with their protein/lipid-binding partners will be an important contribution towards understanding their functional role in the cell. The PXA and PXC domains from SNX-RGS proteins are conserved in SNX14 homologues and appear to be unique to the family, with no recognisable structural or sequence homologues. The Mdm1p PXA domain from yeast cells was found to interact with free fatty acids using TLC (Mike Henne, *Personal communication*), whereas proteomic studies also revealed that the PXA domain interacts with the fatty acid-CoA synthetase, Faa1 (60).

In this chapter, I have described the expression and purification trials of various PXA and PXC domains from several species in both bacterial and insect cell expression systems. The His-tagged PXA domain from *O. latipes* was eventually optimized to produce soluble and folded proteins when expressed with low concentrations of oleic acid and purified with low concentrations of β-octylglucoside detergent. Although I recently obtained micro-needle crystals of PXA domain, I have not yet been able to produce macro-crystals from these conditions. Therefore the structural studies of the PXA domain will still require optimisation in the future. If successful, these studies will help us to identify binding sites for proteins and/or lipids, thus allowing a molecular characterisation of its functional

importance in the cell.

The interaction of ER with the endolysosomal system is rapidly emerging as an area of immense interest in the field of cell biology, but the specific protein machineries that mediate these interactions remain poorly characterized. As the SNX14-related proteins are themselves emerging as key players in this process, and are also involved in human neurological disorders, further molecular insights will be important to allow us to understand how this domain mediates membrane tethering, and other processes such as lipid metabolism, lipophagy and autophagy.

Chapter 8

Chapter 8

Summary and Future Directions

Although all sorting nexin (SNX) proteins fall under the same larger protein family, the PX-proteins, their diversity of structures and functions mean it is impossible to provide a single succinct description of their cellular roles. Over the past decade there has been a great deal of progress in understanding how the Phox-homology (PX) domain of SNX proteins mediates their membrane-localization and recruitment to specific cellular organelles. Although we now have a reasonable structural understanding of how the PX domain interacts with lipids, in particular the endosomal phosphoinositide phosphatidylinositol-3-phosphate (PtdIns3*P*), there is a lack of information regarding how the PX domain regulates various protein–protein interactions. It is crucial to combine both the cellular and structural insights into the molecular functions of the PX proteins to gain a better understanding of how they regulate cellular trafficking and signaling. There is also a lack of knowledge regarding the structures of many of the associated domains within the PX proteins, outside the PX domain itself.

In this thesis, I have studied two subfamilies of the PX-proteins; the SNX-BAR and the SNX-RGS subfamilies. In the first part of my thesis (**Chapters 2-5**), my work focused on the role of PX domains of the distinct SNX-BAR proteins (SNX5, SNX6 and SNX32) in pathogen invasion, and provided the very first look into the structural mechanism of how the PX domains interact with another protein. These chapters provide information regarding how PX domains regulate protein-protein interactions and also demonstrate the potential functional role of SNX-BAR PX domains as a site for the normal transporting of proteins/cargoes that is fundamental to all living cells. In the second part of my thesis (**Chapters 6-7**), my studies are focused on the structural and functional characterization of domains other than the PX-domain in the SNX-RGS subfamily, with a major focus on the RGS domain and the structurally uncharacterised PXA domain. The RGS domain potentially acts as GAP proteins for $G\alpha_s$ subunits, whereas the PXA domain has not only been shown to have the ability to bind to both proteins and lipids but could also control membrane tethering and regulate lipid transfer analogous to other proteins such as the ORP family members (Mesmin et al., 2013). This domain may also play a role in regulation of lipid droplet biogenesis at the ER-lipid droplet contact site.

Chapter 6 reports my findings about the GTPase binding and GAP activity of the

RGS domains of the three human SNX-RGS proteins, which tells us a confirming story of some aspects of previously published studies (48), but questioning others. **Chapter 7** details the extensive efforts I made to optimise the expression and purification of a number of recombinant PXA domains from various homologues and species (**Chapter 7, Table 7**).

8.1 The first insights into how PX domains regulate protein-protein interactions

Chlamydia trachomatis, commonly known as chlamydia, must enter the human host cell to replicate and eventually cause infection. Once inside the host cell, the bacteria develop certain membrane-bound structures called inclusions, where they grow and multiply. As the infection progresses, this inclusion membrane gets modified with many bacterial effector proteins called inclusion membrane proteins. The major role of these inclusion membrane proteins is to hijack host cell proteins and lipids to maintain the inclusion membrane for their own survival within the host cell. SNX5, SNX6 and SNX32 proteins from the SNX-BAR subfamily were previously shown to be hijacked by one such Chlamydial effector protein called Inclusion membrane protein E (IncE) and recruited to the inclusion membrane during Chlamydial infection (26,113). How IncE or any other inclusion membrane proteins hijack human host cell proteins were unknown until now.

Chapter 2 deciphers the structural mechanism through which IncE specifically binds to SNX5, SNX6 and SNX32 proteins. The direct association between the C-terminal region of IncE and the PX domains of SNX5, SNX6 and SNX32 were confirmed using isothermal titration calorimetry. The affinities (K_d) of IncE for SNX5, SNX6 and SNX32 were essentially similar (0.9, 1.1 and 1 uM respectively), which confirms a common recruitment motif within these proteins (**Chapter 2, Fig. 16**), which was further confirmed by recruitment of these proteins to inclusion membranes in infected cells (**Chapter 2, Fig. 15**). The three-dimensional crystal structure of SNX5 PX in complex with the C-terminal region of IncE reveals the IncE forms a long β-hairpin when bound to SNX5 at the base of the extended helical region (**Chapter 2, Fig. 18**), which is a unique structural element found only in SNX5, SNX6 and SNX32 and not in any other SNXs. **Chapter 3** reports the crystal structure of SNX32PX-IncE complex, which confirms the biophysical data from chapter 2. The binding mechanism of SNX32PX-IncE is essentially identical to the SNX5PX-IncE binding mechanism thus further confirming the common binding motif within these proteins required for the IncE binding (**Chapter 3, Fig. 32)**. From the perspective of understanding the endogenous functions of the SNX5, SNX6 and SNX32 proteins, the work in chapter 2 and 3 reveal that the PX domains of these proteins bind IncE using a

site that is both different to the usual lipid binding region common to other PX domains, but is very highly conserved in the protein family. This led me to speculate that IncE might be mimicking and displacing certain important host cell transmembrane proteins/cargo that would normally bind to SNX5, SNX32 and SNX6 (**Fig. 76**). In **Chapter 4** I present important new unpublished data that this is indeed the case, demonstrating the functional role of these specific SNX-BAR proteins as a scaffolds for the binding and normal recycling of cargoes such as the CI-MPR in the cell.

What other potential cargoes might bind to the PX domains of SNX5, SNX6 and SNX32? **Chapter 4** reports initial answers to this question. The Vps retromer core (Vps35, Vps29 and Vps26) and SNX-BAR dimers (SNX1 or SNX2 heterodimerised with SNX5 or SNX6) have long been considered the core retromer components in mammalian cells. Even though no proper biochemical evidence of a stable interaction between the Vps core and SNX-BAR sub complexes in mammalian cells are reported, their mutual functional role in the retrograde trafficking of of cation-independent mannose 6-phosphate (CI-MPR) was accepted as a connection between these two sub complexes. This cooperation of subcomplexes and thus the model of retromer activity were recently questioned by two different labs (115,116), where they found SNX-BAR proteins play a key role in CI-MPR cargo recruitment that appears to be independent of the retromer Vps core. Besides CI-MPR, proteomic studies have revealed many cargos that depend on SNX5, SNX6 and SNX32 for their recycling, including IGF1R and SEMA4C. In collaboration with Dr. Boris Simonetti and Prof. Pete Cullen (University of Bristol, UK), the potential regions on the cytoplasmic tails of these cargoes that bind to SNX5 were mapped. In **Chapter 4**, using isothermal titration calorimetry, the PX domains of SNX5, SNX6 and SNX32 were confirmed to directly bind to the cytoplasmic tails from CIMPR, IGF1R and SEMA4C. The first crystal structure of a PX domain of a SNX5 in complex with a transmembrane cargo, CIMPR is also reported in this chapter, which confirms the binding site on SNX5PX as the same binding site recognised by IncE thus confirming my hypothesis from chapter 2 and 3 that the PX domains from SNX5, SNX6 and SNX32 proteins act as a site for trafficking of cargos. This study also validates the hypothesis from the recent publications that SNX-BAR proteins can bind directly to cargos and function independent of the Vps core subcomplex. Since the crystal structure data of SNX5PX-CIMPR is of low resolution, optimisation of native and seleno-methione crystals of SNX5PX-CIMPR complex are underway, which will enable me to understand the interactions and to perform structure-based mutagenesis to confirm the cargo interaction mechanisms *in vitro*, and to test for defective cargo trafficking *in vivo*. In the future, crystal structures of SNX5PX with IGF1R

and SEMA4C, besides CIMPR, will enable us to understand the different conformations these cargoes adopt to bind to SNX5PX, as well as provide insights into how SNX-BAR protein can allow the trafficking of these cargos to different destinations. Since retromer-SNX cargo sorting pathways has important implications in neurological diseases, a complete understanding of these selected endosomal sorting pathways are likely to open up new opportunities for the design of therapeutic interventions in the future.

As I discussed in chapter 2 and 3, IncE binds to SNX-BAR proteins as a β-hairpin consisting of two antiparallel β-strands, which are connected through a β-turn. Since the two IncE β-strands are positioned close together I developed an idea that the N and C termini of IncE could be linked to form a cyclic peptide, which might demonstrate increased binding for SNX-BAR proteins and stability when compared to the linear IncE peptide, due to the increased rigidity and decreased entropy. **Chapter 5** explores this idea of a more stable cyclic IncE peptide, which in the future, could be used as a specific inhibitory tool to study the role of SNX5 in membrane trafficking. The work reported in this chapter was performed in collaboration with Prof. David Fairlie and Dr. Timothy Hill (IMB, University of Queensland). In this chapter, I have identified a cyclic IncE peptide that has the ability to bind to SNX5 PX domain with an increased binding affinity, which is nearly 30x higher than the standard IncE peptide, and ~300x higher than cellular cargo molecules. This suggests a strong potential to inhibit the binding of a transmembrane cargo such as CIMPR, confirmed *in vitro*, and the next steps will be to determine the ability of such peptides to cross the membrane so that they can be used as tools for studying the trafficking of SNX-bound cargos. Further into the future such peptides may have additional uses such as testing the importance of such pathways in disease states and their therapeutic targeting potential.

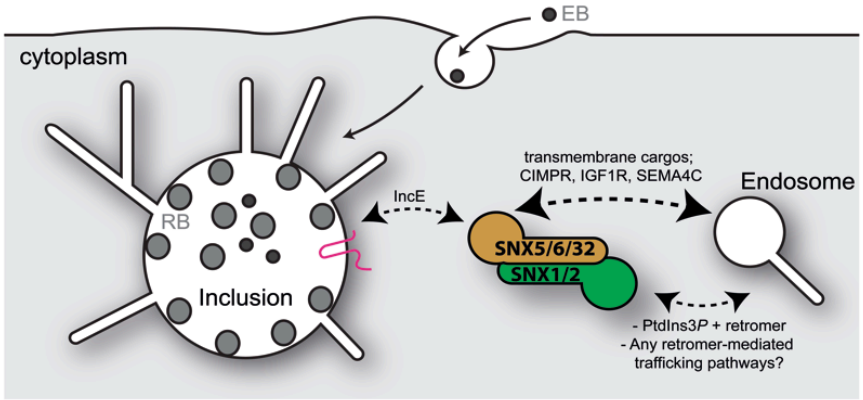


Figure 76. *Schematic representation of the involvement of PX domains of SNX5, SNX6 and SNX32 in interaction with IncE which mimicks the transmembrane cargos CIMPR, IGF1R and SEMA4C that normally binds to SNX-BAR proteins for their proper recycling.*

8.2 Understanding the various functions of different domains in the SNX-RGS subfamily of proteins and insights into how they might be working together in our cell.

In studies of a different sub-family of SNX proteins, **Chapter 6** explores the ability of the RGS domains of SNX-RGS proteins (SNX13, SNX14 and SNX25) to bind to $G\alpha_s$ and their ability to hydrolyze GTP and attenuate $G\alpha_s$ signaling. Even though my studies demonstrated the ability of SNX-RGS proteins to successfully interact with $G\alpha_s$, none of them were found to possess substantial GAP activity. This result is in contrast to the original studies of SNX13-RGS, which identified a potential $G\alpha_s$-GAP activity (48), whereas it is in support to the results of Ha *et al.* (49), which failed to identify a GAP activity for SNX14. This leads me to the speculate that since SNX-RGS proteins bind to $G\alpha_s$, these proteins may still act as negative regulators of $G\alpha_s$ signaling, whereby they will sequester the G-protein and prevent its interactions with downstream effectors. Alternatively, the SNX-RGS proteins may not be regulators of $G\alpha_s$, but may in themselves act as specific effectors of these signaling molecules. My studies (**chapter 6**) also reported SNX-RGS proteins having different affinities for different $G\alpha_s$-nucleotide complexes. They exhibited a preference for either the transition state conformation or the active state of the $G\alpha_s$ protein. In contrast to the previous reports by Zheng *et al*. (48), SNX13 and SNX14 RGS proteins also demonstrated weaker interaction with the inactive state of the molecule; $G\alpha_s$-GDP. This result led me to conclude that hydrolysis of GTP by $G\alpha_s$ may not be enough to fully terminate SNX-RGS binding, and/or the SNX-RGS might play a role in suppressing the signaling of $G\alpha_s$ signaling by also sequestering $G\alpha_s$-GDP. The preference of RGS domains for one $G\alpha_s$–nucleotide bound state over the other also suggests that the binding sites of the RGS proteins adopts different favourable conformations to bind the preferred $G\alpha_s$ state. Most RGS domain interactions are typically in the micromolar range similar to what I saw in my experiments, which means these interactions are quite likely to occur physiologically. Even though my preliminary studies demonstrate the RGS domains of SNX-RGS do not possess GAP activity for $G\alpha_s$, both SNX13 and SNX14 proteins are previously reported to regulate cAMP signaling (48,49). Whether RGS domains from SNX-RGS proteins share pathways of feedback phosphorylation or palmitoylation for the regulation of cAMP signalling like other classical RGS proteins, are yet to be explored. In the future, the crystal structures of the RGS domains in complex with $G\alpha_s$ will enable us to understand the residues providing specificity between these partners and how they achieve preference for the different $G\alpha_s$-nucleotide states. The longer-term goal is to

define the physiological purpose of SNX-RGS proteins in regulating $G\alpha_s$ mediated cAMP signaling. An open question is also where these $G\alpha_s$ interactions occur in the cell. As described in past work by the Collins lab and others, the various SNX-RGS proteins appear to be predominantly associated with internal organelles including the ER and endosomes, as well as lipid droplet contact sites. Thus the role of SNX-RGS interactions with $G\alpha_s$ is likely to be specifically associated with intracellular compartments distinct from the site of GPCR-$G\alpha$ activation at the cell surface.

Exploring the structure of the uncharacterised PXA domain of SNX-RGS proteins will enable us to understand the molecular mechanisms by which this domain binds to lipids and proteins, and reveal how they contribute towards their function in regulation of lipid droplet biogenesis at the ER-LD contact site.

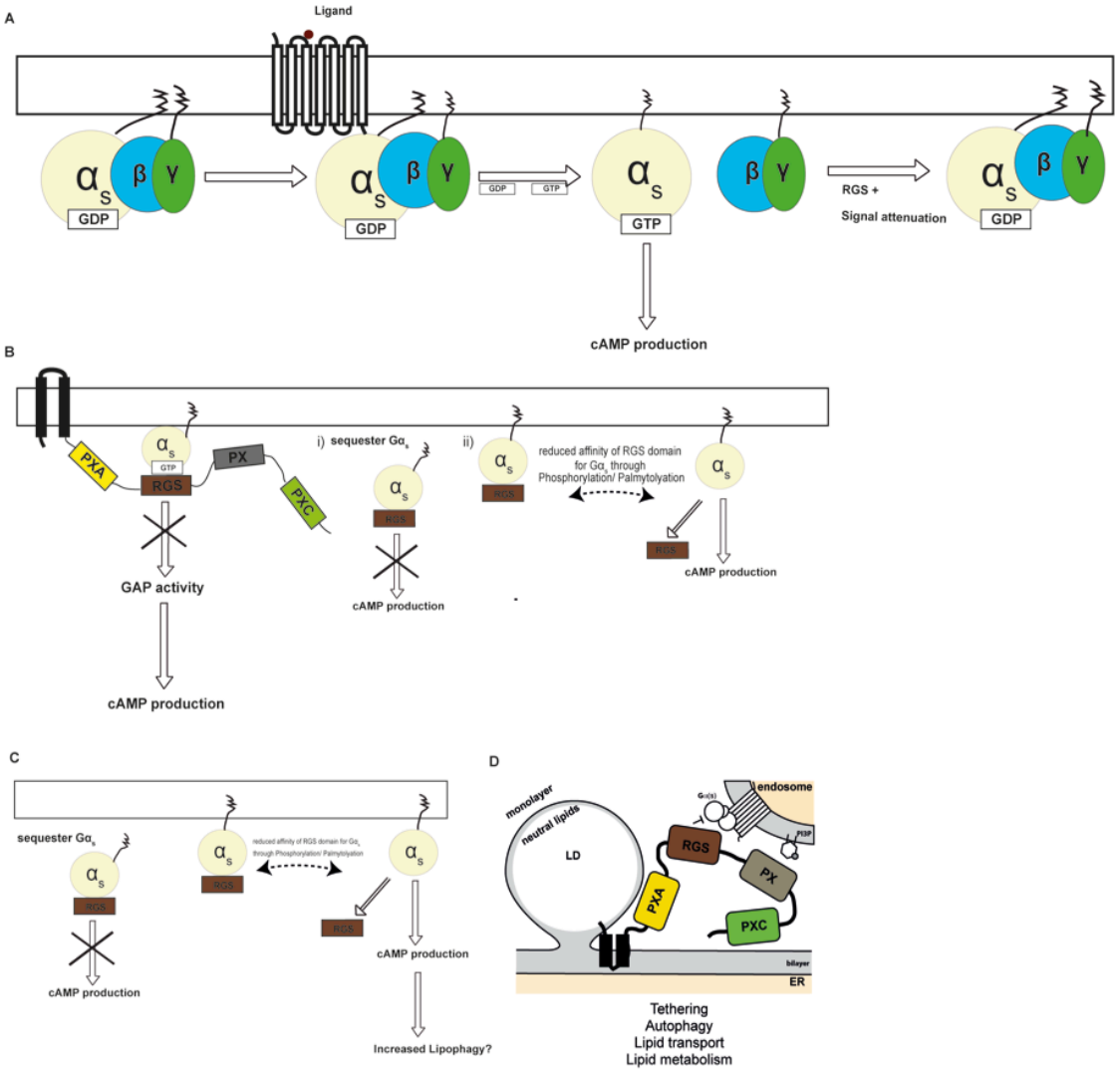


***Figure 77**. A) RGS proteins with GAP activity. B) RGS domains of SNX-RGS might act as negative regulators of $G\alpha_s$ signaling by sequestering the G-protein, whereas the increased cAMP production can be regulated by either palmytoylation or feedback phosphorylation of RGS and thus inhibiting its affinity for $G\alpha_s$. C) Could RGS domain be involved in lipophagy by regulating cAMP*

production? D) Schematic representation of the involvement of SNX-RGS protein in lipid droplet biogenesis.

Chapter 7 describes in detail the efforts to successfully express and purify various PXA domains from a wide variety of species in both bacterial and insect cell expression systems. Even though this chapter presents many difficulties while trying to express and purify these domains, I report how I eventually optimized the conditions to successfully produce soluble and folded PXA domain protein. Out of the various constructs from different species that I tried, the His-tagged PXA domain constructs from *Oryzias latipes* (**Chapter 7, Table 10**) (in collaboration with Assoc. Prof. Mike Henne, University of Texas) lead to the production of folded protein, which in turn produced micro needle crystals. The condition that proved to be efficient for the successful expression and purification was the expression of these constructs with low concentrations of oleic acid and purification with low concentrations of β-octylglucoside detergent. Future efforts will now hopefully extend these preliminary studies to produce diffraction quality crystals suitable for structure determination, which will provide important insights into their so far unknown structures and functions.

8.3 Concluding remarks

Overall, the research presented in this thesis describes the first protein-protein structural analysis of a PX domain of the SNX subfamily, which is usually known for its PI3*P* binding abilities, and how domains other than the PX domain work for the normal functioning of the SNX proteins in the cell. My work has provided new insights into how the PX domains regulate protein-protein interactions, how they contribute to pathogen invasion and different diseases, and should inform future studies to determine whether they might be suitable as therapeutic targets and if so, how such interventions can be achieved.

References

1. Alan, R. (1991) Glossary of genetics (classical and molecular) fifth edition: By R Rieger, A Michaelis and M M Green. pp 553. Springer-Verlag, Berlin, 1991. DM69 ISBN 0-540-52054-6. *Biochemical Education* **19**, 177-177
2. Mayor, S., Presley, J. F., and Maxfield, F. R. (1993) Sorting of membrane components from endosomes and subsequent recycling to the cell surface occurs by a bulk flow process. *J Cell Biol* **121**, 1257-1269
3. Jovic, M., Sharma, M., Rahajeng, J., and Caplan, S. (2010) The early endosome: a busy sorting station for proteins at the crossroads. *Histol Histopathol* **25**, 99-112
4. Hsu, V. W., Bai, M., and Li, J. (2012) Getting active: protein sorting in endocytic recycling. *Nat Rev Mol Cell Biol* **13**, 323-328
5. Worby, C. A., and Dixon, J. E. (2002) Sorting out the cellular functions of sorting nexins. *Nat Rev Mol Cell Biol* **3**, 919-931
6. Beglova, N., and Blacklow, S. C. (2005) The LDL receptor: how acid pulls the trigger. *Trends Biochem Sci* **30**, 309-317
7. Cullen, P. J. (2008) Endosomal sorting and signalling: an emerging role for sorting nexins. *Nat Rev Mol Cell Biol* **9**, 574-582
8. Teasdale, R. D., and Collins, B. M. (2012) Insights into the PX (phox-homology) domain and SNX (sorting nexin) protein families: structures, functions and roles in disease. *Biochem J* **441**, 39-59
9. van Weering, J. R., Verkade, P., and Cullen, P. J. (2010) SNX-BAR proteins in phosphoinositide-mediated, tubular-based endosomal sorting. *Semin Cell Dev Biol* **21**, 371-380
10. Klinger, S. C., Siupka, P., and Nielsen, M. S. (2015) Retromer-Mediated Trafficking of Transmembrane Receptors and Transporters. *Membranes (Basel)* **5**, 288-306
11. van Weering, J. R., and Cullen, P. J. (2014) Membrane-associated cargo recycling by tubule-based endosomal sorting. *Semin Cell Dev Biol* **31**, 40-47
12. Bonifacino, J. S., and Rojas, R. (2006) Retrograde transport from endosomes to the trans-Golgi network. *Nat Rev Mol Cell Biol* **7**, 568-579
13. Johannes, L., and Popoff, V. (2008) Tracing the retrograde route in protein trafficking. *Cell* **135**, 1175-1187
14. Ghosh, P., Dahms, N. M., and Kornfeld, S. (2003) Mannose 6-phosphate receptors: new twists in the tale. *Nat Rev Mol Cell Biol* **4**, 202-212

15. Bonifacino, J. S., and Hurley, J. H. (2008) Retromer. *Curr Opin Cell Biol* **20**, 427-436
16. Gallon, M., and Cullen, P. J. (2015) Retromer and sorting nexins in endosomal sorting. *Biochem Soc Trans* **43**, 33-47
17. Carlton, J., Bujny, M., Peter, B. J., Oorschot, V. M., Rutherford, A., Mellor, H., Klumperman, J., McMahon, H. T., and Cullen, P. J. (2004) Sorting nexin-1 mediates tubular endosome-to-TGN transport through coincidence sensing of high- curvature membranes and 3-phosphoinositides. *Curr Biol* **14**, 1791-1800
18. Wassmer, T., Attar, N., Bujny, M. V., Oakley, J., Traer, C. J., and Cullen, P. J. (2007) A loss-of-function screen reveals SNX5 and SNX6 as potential components of the mammalian retromer. *J Cell Sci* **120**, 45-54
19. Small, S. A., Kent, K., Pierce, A., Leung, C., Kang, M. S., Okada, H., Honig, L., Vonsattel, J. P., and Kim, T. W. (2005) Model-guided microarray implicates the retromer complex in Alzheimer's disease. *Ann Neurol* **58**, 909-919
20. Vilarino-Guell, C., Wider, C., Ross, O. A., Dachsel, J. C., Kachergus, J. M., Lincoln, S. J., Soto-Ortolaza, A. I., Cobb, S. A., Wilhoite, G. J., Bacon, J. A., Behrouz, B., Melrose, H. L., Hentati, E., Puschmann, A., Evans, D. M., Conibear, E., Wasserman, W. W., Aasly, J. O., Burkhard, P. R., Djaldetti, R., Ghika, J., Hentati, F., Krygowska-Wajs, A., Lynch, T., Melamed, E., Rajput, A., Rajput, A. H., Solida, A., Wu, R. M., Uitti, R. J., Wszolek, Z. K., Vingerhoets, F., and Farrer, M. J. (2011) VPS35 mutations in Parkinson disease. *Am J Hum Genet* **89**, 162-167
21. Zimprich, A., Benet-Pages, A., Struhal, W., Graf, E., Eck, S. H., Offman, M. N., Haubenberger, D., Spielberger, S., Schulte, E. C., Lichtner, P., Rossle, S. C., Klopp, N., Wolf, E., Seppi, K., Pirker, W., Presslauer, S., Mollenhauer, B., Katzenschlager, R., Foki, T., Hotzy, C., Reinthaler, E., Harutyunyan, A., Kralovics, R., Peters, A., Zimprich, F., Brucke, T., Poewe, W., Auff, E., Trenkwalder, C., Rost, B., Ransmayr, G., Winkelmann, J., Meitinger, T., and Strom, T. M. (2011) A mutation in VPS35, encoding a subunit of the retromer complex, causes late-onset Parkinson disease. *Am J Hum Genet* **89**, 168-175
22. Rovelet-Lecrux, A., Charbonnier, C., Wallon, D., Nicolas, G., Seaman, M. N., Pottier, C., Breusegem, S. Y., Mathur, P. P., Jenardhanan, P., Le Guennec, K., Mukadam, A. S., Quenez, O., Coutant, S., Rousseau, S., Richard, A. C., Boland, A., Deleuze, J. F., Frebourg, T., Hannequin, D., Campion, D., and collaborators, C.-M. (2015) De novo deleterious genetic variations target a biological network centered on Abeta peptide in early-onset Alzheimer disease. *Mol Psychiatry* **20**, 1046-1056

23. Pylypenko, O., Lundmark, R., Rasmuson, E., Carlsson, S. R., and Rak, A. (2007) The PX-BAR membrane-remodeling unit of sorting nexin 9. *EMBO J* **26**, 4788-4800
24. Bravo, J., Karathanassis, D., Pacold, C. M., Pacold, M. E., Ellson, C. D., Anderson, K. E., Butler, P. J., Lavenir, I., Perisic, O., Hawkins, P. T., Stephens, L., and Williams, R. L. (2001) The crystal structure of the PX domain from p40(phox) bound to phosphatidylinositol 3-phosphate. *Mol Cell* **8**, 829-839
25. Koharudin, L. M., Furey, W., Liu, H., Liu, Y. J., and Gronenborn, A. M. (2009) The phox domain of sorting nexin 5 lacks phosphatidylinositol 3-phosphate (PtdIns(3)P) specificity and preferentially binds to phosphatidylinositol 4,5-bisphosphate (PtdIns(4,5)P2). *J Biol Chem* **284**, 23697-23707
26. Mirrashidi, K. M., Elwell, C. A., Verschueren, E., Johnson, J. R., Frando, A., Von Dollen, J., Rosenberg, O., Gulbahce, N., Jang, G., Johnson, T., Jager, S., Gopalakrishnan, A. M., Sherry, J., Dunn, J. D., Olive, A., Penn, B., Shales, M., Cox, J. S., Starnbach, M. N., Derre, I., Valdivia, R., Krogan, N. J., and Engel, J. (2015) Global Mapping of the Inc-Human Interactome Reveals that Retromer Restricts Chlamydia Infection. *Cell Host Microbe* **18**, 109-121
27. Aeberhard, L., Banhart, S., Fischer, M., Jehmlich, N., Rose, L., Koch, S., Laue, M., Renard, B. Y., Schmidt, F., and Heuer, D. (2015) The Proteome of the Isolated Chlamydia trachomatis Containing Vacuole Reveals a Complex Trafficking Platform Enriched for Retromer Components. *PLoS Pathog* **11**, e1004883
28. Schachter, J., Storz, J., Tarizzo, M. L., and Bogel, K. (1973) Chlamydiae as agents of human and animal diseases. *Bull World Health Organ* **49**, 443-449
29. Moulder, J. W. (1991) Interaction of chlamydiae and host cells in vitro. *Microbiol Rev* **55**, 143-190
30. Hybiske, K. (2015) Expanding the Molecular Toolkit for Chlamydia. *Cell Host Microbe* **18**, 11-13
31. Moore, E. R., and Ouellette, S. P. (2014) Reconceptualizing the chlamydial inclusion as a pathogen-specified parasitic organelle: an expanded role for Inc proteins. *Frontiers in cellular and infection microbiology* **4**, 157
32. Lutter, E. I., Martens, C., and Hackstadt, T. (2012) Evolution and conservation of predicted inclusion membrane proteins in chlamydiae. *Comp Funct Genomics* **2012**, 362104
33. Heuer, D., Rejman Lipinski, A., Machuy, N., Karlas, A., Wehrens, A., Siedler, F., Brinkmann, V., and Meyer, T. F. (2009) Chlamydia causes fragmentation of the Golgi compartment to ensure reproduction. *Nature* **457**, 731-735

34. Frost, A., Unger, V. M., and De Camilli, P. (2009) The BAR domain superfamily: membrane-molding macromolecules. *Cell* **137**, 191-196
35. Groppelli, E., Len, A. C., Granger, L. A., and Jolly, C. (2014) Retromer regulates HIV-1 envelope glycoprotein trafficking and incorporation into virions. *PLoS Pathog* **10**, e1004518
36. Lipovsky, A., Popa, A., Pimienta, G., Wyler, M., Bhan, A., Kuruvilla, L., Guie, M. A., Poffenberger, A. C., Nelson, C. D., Atwood, W. J., and DiMaio, D. (2013) Genome-wide siRNA screen identifies the retromer as a cellular entry factor for human papillomavirus. *Proc Natl Acad Sci U S A* **110**, 7452-7457
37. McDonough, J. A., Newton, H. J., Klum, S., Swiss, R., Agaisse, H., and Roy, C. R. (2013) Host pathways important for Coxiella burnetii infection revealed by genome-wide RNA interference screening. *MBio* **4**, e00606-00612
38. Mas, C., Norwood, S. J., Bugarcic, A., Kinna, G., Leneva, N., Kovtun, O., Ghai, R., Ona Yanez, L. E., Davis, J. L., Teasdale, R. D., and Collins, B. M. (2014) Structural basis for different phosphoinositide specificities of the PX domains of sorting nexins regulating G-protein signaling. *J Biol Chem* **289**, 28554-28568
39. Gilman, A. G. (1987) G proteins: transducers of receptor-generated signals. *Annu Rev Biochem* **56**, 615-649
40. Johnston, C. A., and Siderovski, D. P. (2007) Receptor-mediated activation of heterotrimeric G-proteins: current structural insights. *Mol Pharmacol* **72**, 219-230
41. Zheng, B., Lavoie, C., Tang, T. D., Ma, P., Meerloo, T., Beas, A., and Farquhar, M. G. (2004) Regulation of epidermal growth factor receptor degradation by heterotrimeric Galphas protein. *Mol Biol Cell* **15**, 5538-5550
42. Ross, E. M., and Wilkie, T. M. (2000) GTPase-activating proteins for heterotrimeric G proteins: regulators of G protein signaling (RGS) and RGS-like proteins. *Annu Rev Biochem* **69**, 795-827
43. Stewart, A., Huang, J., and Fisher, R. A. (2012) RGS Proteins in Heart: Brakes on the Vagus. *Front Physiol* **3**, 95
44. Coleman, D. E., Berghuis, A. M., Lee, E., Linder, M. E., Gilman, A. G., and Sprang, S. R. (1994) Structures of active conformations of Gi alpha 1 and the mechanism of GTP hydrolysis. *Science* **265**, 1405-1412
45. Kimple, A. J., Bosch, D. E., Giguere, P. M., and Siderovski, D. P. (2011) Regulators of G-protein signaling and their Galpha substrates: promises and challenges in their use as drug discovery targets. *Pharmacol Rev* **63**, 728-749

46. Soundararajan, M., Willard, F. S., Kimple, A. J., Turnbull, A. P., Ball, L. J., Schoch, G. A., Gileadi, C., Fedorov, O. Y., Dowler, E. F., Higman, V. A., Hutsell, S. Q., Sundstrom, M., Doyle, D. A., and Siderovski, D. P. (2008) Structural diversity in the RGS domain and its interaction with heterotrimeric G protein alpha-subunits. *Proc Natl Acad Sci U S A* **105**, 6457-6462

47. Tesmer, J. J., Berman, D. M., Gilman, A. G., and Sprang, S. R. (1997) Structure of RGS4 bound to AlF4--activated G(i alpha1): stabilization of the transition state for GTP hydrolysis. *Cell* **89**, 251-261

48. Zheng, B., Ma, Y. C., Ostrom, R. S., Lavoie, C., Gill, G. N., Insel, P. A., Huang, X. Y., and Farquhar, M. G. (2001) RGS-PX1, a GAP for GalphaS and sorting nexin in vesicular trafficking. *Science* **294**, 1939-1942

49. Ha, C. M., Park, D., Kim, Y., Na, M., Panda, S., Won, S., Kim, H., Ryu, H., Park, Z. Y., Rasenick, M. M., and Chang, S. (2015) SNX14 is a bifunctional negative regulator for neuronal 5-HT6 receptor signaling. *J Cell Sci* **128**, 1848-1861

50. Li, J., Li, C., Zhang, D., Shi, D., Qi, M., Feng, J., Yuan, T., Xu, X., Liang, D., Xu, L., Zhang, H., Liu, Y., Chen, J., Ye, J., Jiang, W., Cui, Y., Zhang, Y., Peng, L., Zhou, Z., and Chen, Y. H. (2014) SNX13 reduction mediates heart failure through degradative sorting of apoptosis repressor with caspase recruitment domain. *Nat Commun* **5**, 5177

51. Hao, X., Wang, Y., Ren, F., Zhu, S., Ren, Y., Jia, B., Li, Y. P., Shi, Y., and Chang, Z. (2011) SNX25 regulates TGF-beta signaling by enhancing the receptor degradation. *Cellular signalling* **23**, 935-946

52. Du, Y., Zou, Y., Yu, W., Shi, R., Zhang, M., Yang, W., Duan, J., Deng, Y., Wang, X., and Lu, Y. (2013) Expression pattern of sorting Nexin 25 in temporal lobe epilepsy: a study on patients and pilocarpine-induced rats. *Brain Res* **1509**, 79-85

53. Suh, J. M., Stenesen, D., Peters, J. M., Inoue, A., Cade, A., and Graff, J. M. (2008) An RGS-containing sorting nexin controls Drosophila lifespan. *PLoS One* **3**, e2152

54. Akizu, N., Cantagrel, V., Zaki, M. S., Al-Gazali, L., Wang, X., Rosti, R. O., Dikoglu, E., Gelot, A. B., Rosti, B., Vaux, K. K., Scott, E. M., Silhavy, J. L., Schroth, J., Copeland, B., Schaffer, A. E., Gordts, P. L., Esko, J. D., Buschman, M. D., Field, S. J., Napolitano, G., Abdel-Salam, G. M., Ozgul, R. K., Sagiroglu, M. S., Azam, M., Ismail, S., Aglan, M., Selim, L., Mahmoud, I. G., Abdel-Hadi, S., Badawy, A. E., Sadek, A. A., Mojahedi, F., Kayserili, H., Masri, A., Bastaki, L., Temtamy, S., Muller, U., Desguerre, I., Casanova, J. L., Dursun, A., Gunel, M., Gabriel, S. B., de Lonlay, P., and Gleeson, J. G. (2015) Biallelic mutations in SNX14 cause a syndromic form

of cerebellar atrophy and lysosome-autophagosome dysfunction. *Nat Genet* **47**, 528-534

55. Thomas, A. C., Williams, H., Seto-Salvia, N., Bacchelli, C., Jenkins, D., O'Sullivan, M., Mengrelis, K., Ishida, M., Ocaka, L., Chanudet, E., James, C., Lescai, F., Anderson, G., Morrogh, D., Ryten, M., Duncan, A. J., Pai, Y. J., Saraiva, J. M., Ramos, F., Farren, B., Saunders, D., Vernay, B., Gissen, P., Straatmaan-Iwanowska, A., Baas, F., Wood, N. W., Hersheson, J., Houlden, H., Hurst, J., Scott, R., Bitner-Glindzicz, M., Moore, G. E., Sousa, S. B., and Stanier, P. (2014) Mutations in SNX14 cause a distinctive autosomal-recessive cerebellar ataxia and intellectual disability syndrome. *Am J Hum Genet* **95**, 611-621
56. Karaca, E., Harel, T., Pehlivan, D., Jhangiani, S. N., Gambin, T., Coban Akdemir, Z., Gonzaga-Jauregui, C., Erdin, S., Bayram, Y., Campbell, I. M., Hunter, J. V., Atik, M. M., Van Esch, H., Yuan, B., Wiszniewski, W., Isikay, S., Yesil, G., Yuregir, O. O., Tug Bozdogan, S., Aslan, H., Aydin, H., Tos, T., Aksoy, A., De Vivo, D. C., Jain, P., Geckinli, B. B., Sezer, O., Gul, D., Durmaz, B., Cogulu, O., Ozkinay, F., Topcu, V., Candan, S., Cebi, A. H., Ikbal, M., Yilmaz Gulec, E., Gezdirici, A., Koparir, E., Ekici, F., Coskun, S., Cicek, S., Karaer, K., Koparir, A., Duz, M. B., Kirat, E., Fenercioglu, E., Ulucan, H., Seven, M., Guran, T., Elcioglu, N., Yildirim, M. S., Aktas, D., Alikasifoglu, M., Ture, M., Yakut, T., Overton, J. D., Yuksel, A., Ozen, M., Muzny, D. M., Adams, D. R., Boerwinkle, E., Chung, W. K., Gibbs, R. A., and Lupski, J. R. (2015) Genes that Affect Brain Structure and Function Identified by Rare Variant Analyses of Mendelian Neurologic Disease. *Neuron* **88**, 499-513
57. Lee, J. H., Cheng, R., Vardarajan, B., Lantigua, R., Reyes-Dumeyer, D., Ortmann, W., Graham, R. R., Bhangale, T., Behrens, T. W., Medrano, M., Jimenez-Velazquez, I. Z., and Mayeux, R. (2015) Genetic Modifiers of Age at Onset in Carriers of the G206A Mutation in PSEN1 With Familial Alzheimer Disease Among Caribbean Hispanics. *JAMA neurology* **72**, 1043-1051
58. Koseki, T., Inohara, N., Chen, S., and Nunez, G. (1998) ARC, an inhibitor of apoptosis expressed in skeletal muscle and heart that interacts selectively with caspases. *Proc Natl Acad Sci U S A* **95**, 5156-5160
59. Henne, W. M., Zhu, L., Balogi, Z., Stefan, C., Pleiss, J. A., and Emr, S. D. (2015) Mdm1/Snx13 is a novel ER-endolysosomal interorganelle tethering protein. *J Cell Biol* **210**, 541-551

60. Hariri, H., Rogers, S., Ugrankar, R., Liu, Y. L., Feathers, J. R., and Henne, W. M. (2018) Lipid droplet biogenesis is spatially coordinated at ER-vacuole contacts under nutritional stress. *EMBO Rep* **19**, 57-72
61. Kivalo, E., and Stjernvall, L. (1958) Vacuolized lymphocytes in juvenile amaurotic idiocy; an electron microscopic study. *Ann Paediatr Fenn* **4**, 25-29
62. Di Russo Case, E., and Samuel, J. E. (2016) Contrasting Lifestyles Within the Host Cell. *Microbiol Spectr* **4**
63. Personnic, N., Barlocher, K., Finsel, I., and Hilbi, H. (2016) Subversion of Retrograde Trafficking by Translocated Pathogen Effectors. *Trends Microbiol* **24**, 450-462
64. Bastidas, R. J., Elwell, C. A., Engel, J. N., and Valdivia, R. H. (2013) Chlamydial intracellular survival strategies. *Cold Spring Harb Perspect Med* **3**, a010256
65. Derre, I. (2015) Chlamydiae interaction with the endoplasmic reticulum: contact, function and consequences. *Cellular microbiology* **17**, 959-966
66. Elwell, C., Mirrashidi, K., and Engel, J. (2016) Chlamydia cell biology and pathogenesis. *Nat Rev Microbiol* **14**, 385-400
67. Aral, S. O., Over, M., Manhart, L., and Holmes, K. K. (2006) Sexually Transmitted Infections. in *Disease Control Priorities in Developing Countries* (Jamison, D. T., Breman, J. G., Measham, A. R., Alleyne, G., Claeson, M., Evans, D. B., Jha, P., Mills, A., and Musgrove, P. eds.), 2nd Ed., Washington (DC). pp
68. Newman, L., Rowley, J., Vander Hoorn, S., Wijesooriya, N. S., Unemo, M., Low, N., Stevens, G., Gottlieb, S., Kiarie, J., and Temmerman, M. (2015) Global Estimates of the Prevalence and Incidence of Four Curable Sexually Transmitted Infections in 2012 Based on Systematic Review and Global Reporting. *PLoS One* **10**, e0143304
69. Mpiga, P., and Ravaoarinoro, M. (2006) Chlamydia trachomatis persistence: an update. *Microbiological research* **161**, 9-19
70. Kohlhoff, S. A., and Hammerschlag, M. R. (2015) Treatment of Chlamydial infections: 2014 update. *Expert opinion on pharmacotherapy* **16**, 205-212
71. Ward, M. E. (1983) Chlamydial classification, development and structure. *Br Med Bull* **39**, 109-115
72. Rockey, D. D., Scidmore, M. A., Bannantine, J. P., and Brown, W. J. (2002) Proteins in the chlamydial inclusion membrane. *Microbes and infection / Institut Pasteur* **4**, 333-340

73. Dehoux, P., Flores, R., Dauga, C., Zhong, G., and Subtil, A. (2011) Multi-genome identification and characterization of chlamydiae-specific type III secretion substrates: the Inc proteins. *BMC genomics* **12**, 109
74. Kostriukova, E. S., Lazarev, V. N., and Govorun, V. M. (2008) [Inclusion membrane proteins of Chlamydiaceae]. *Biomeditsinskaia khimiia* **54**, 24-41
75. Li, Z., Chen, C., Chen, D., Wu, Y., Zhong, Y., and Zhong, G. (2008) Characterization of fifty putative inclusion membrane proteins encoded in the Chlamydia trachomatis genome. *Infection and immunity* **76**, 2746-2757
76. Rzomp, K. A., Moorhead, A. R., and Scidmore, M. A. (2006) The GTPase Rab4 interacts with Chlamydia trachomatis inclusion membrane protein CT229. *Infect Immun* **74**, 5362-5373
77. Cortes, C., Rzomp, K. A., Tvinnereim, A., Scidmore, M. A., and Wizel, B. (2007) Chlamydia pneumoniae inclusion membrane protein Cpn0585 interacts with multiple Rab GTPases. *Infection and immunity* **75**, 5586-5596
78. Delevoye, C., Nilges, M., Dehoux, P., Paumet, F., Perrinet, S., Dautry-Varsat, A., and Subtil, A. (2008) SNARE protein mimicry by an intracellular bacterium. *PLoS Pathog* **4**, e1000022
79. Mital, J., Miller, N. J., Fischer, E. R., and Hackstadt, T. (2010) Specific chlamydial inclusion membrane proteins associate with active Src family kinases in microdomains that interact with the host microtubule network. *Cellular microbiology* **12**, 1235-1249
80. Mital, J., Lutter, E. I., Barger, A. C., Dooley, C. A., and Hackstadt, T. (2015) Chlamydia trachomatis inclusion membrane protein CT850 interacts with the dynein light chain DYNLT1 (Tctex1). *Biochemical and biophysical research communications* **462**, 165-170
81. Dumoux, M., Menny, A., Delacour, D., and Hayward, R. D. (2015) A Chlamydia effector recruits CEP170 to reprogram host microtubule organization. *J Cell Sci* **128**, 3420-3434
82. Lutter, E. I., Barger, A. C., Nair, V., and Hackstadt, T. (2013) Chlamydia trachomatis inclusion membrane protein CT228 recruits elements of the myosin phosphatase pathway to regulate release mechanisms. *Cell reports* **3**, 1921-1931
83. Scidmore, M. A., and Hackstadt, T. (2001) Mammalian 14-3-3beta associates with the Chlamydia trachomatis inclusion membrane via its interaction with IncG. *Molecular microbiology* **39**, 1638-1650

84. Kokes, M., Dunn, J. D., Granek, J. A., Nguyen, B. D., Barker, J. R., Valdivia, R. H., and Bastidas, R. J. (2015) Integrating chemical mutagenesis and whole-genome sequencing as a platform for forward and reverse genetic analysis of Chlamydia. *Cell Host Microbe* **17**, 716-725
85. Derre, I., Swiss, R., and Agaisse, H. (2011) The lipid transfer protein CERT interacts with the Chlamydia inclusion protein IncD and participates to ER-Chlamydia inclusion membrane contact sites. *PLoS Pathog* **7**, e1002092
86. Elwell, C. A., Jiang, S., Kim, J. H., Lee, A., Wittmann, T., Hanada, K., Melancon, P., and Engel, J. N. (2011) Chlamydia trachomatis co-opts GBF1 and CERT to acquire host sphingomyelin for distinct roles during intracellular development. *PLoS Pathog* **7**, e1002198
87. van Weering, J. R., Sessions, R. B., Traer, C. J., Kloer, D. P., Bhatia, V. K., Stamou, D., Carlsson, S. R., Hurley, J. H., and Cullen, P. J. (2012) Molecular basis for SNX-BAR-mediated assembly of distinct endosomal sorting tubules. *The EMBO journal* **31**, 4466-4480
88. Huston, W. M., Theodoropoulos, C., Mathews, S. A., and Timms, P. (2008) Chlamydia trachomatis responds to heat shock, penicillin induced persistence, and IFN-gamma persistence by altering levels of the extracytoplasmic stress response protease HtrA. *BMC microbiology* **8**, 190
89. Harris, S. R., Clarke, I. N., Seth-Smith, H. M., Solomon, A. W., Cutcliffe, L. T., Marsh, P., Skilton, R. J., Holland, M. J., Mabey, D., Peeling, R. W., Lewis, D. A., Spratt, B. G., Unemo, M., Persson, K., Bjartling, C., Brunham, R., de Vries, H. J., Morre, S. A., Speksnijder, A., Bebear, C. M., Clerc, M., de Barbeyrac, B., Parkhill, J., and Thomson, N. R. (2012) Whole-genome analysis of diverse Chlamydia trachomatis strains identifies phylogenetic relationships masked by current clinical typing. *Nature genetics* **44**, 413-419, S411
90. Kerr, M. C., Lindsay, M. R., Luetterforst, R., Hamilton, N., Simpson, F., Parton, R. G., Gleeson, P. A., and Teasdale, R. D. (2006) Visualisation of macropinosome maturation by the recruitment of sorting nexins. *J Cell Sci* **119**, 3967-3980
91. Kerr, M. C., Castro, N. A., Karunaratne, S., and Teasdale, R. D. (2012) The Phosphoinositides: Key Regulators of Salmonella Containing Vacuole (SCV) Trafficking and Identity. in *Salmonella - Distribution, Adaptation, Control Measures and Molecular Technologies* (Annous, B. A., and Gurtler, J. B. eds.), InTech. pp

92. Wang, J. T., Kerr, M. C., Karunaratne, S., Jeanes, A., Yap, A. S., and Teasdale, R. D. (2010) The SNX-PX-BAR family in macropinocytosis: the regulation of macropinosome formation by SNX-PX-BAR proteins. *PLoS One* **5**, e13763
93. Battye, T. G., Kontogiannis, L., Johnson, O., Powell, H. R., and Leslie, A. G. (2011) iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr D Biol Crystallogr* **67**, 271-281
94. Evans, P. R., and Murshudov, G. N. (2013) How good are my data and what is the resolution? *Acta crystallographica. Section D, Biological crystallography* **69**, 1204-1214
95. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A., and Wilson, K. S. (2011) Overview of the CCP4 suite and current developments. *Acta crystallographica. Section D, Biological crystallography* **67**, 235-242
96. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *Journal of applied crystallography* **40**, 658-674
97. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta crystallographica. Section D, Biological crystallography* **66**, 486-501
98. Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Echols, N., Headd, J. J., Hung, L. W., Jain, S., Kapral, G. J., Grosse Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R. D., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2011) The Phenix software for automated determination of macromolecular structures. *Methods* **55**, 94-106
99. Teo, W. X., Kerr, M. C., Huston, W. M., and Teasdale, R. D. (2016) Sortilin is associated with the chlamydial inclusion and is modulated during infection. *Biol Open* **5**, 429-435
100. Bolte, S., and Cordelieres, F. P. (2006) A guided tour into subcellular colocalization analysis in light microscopy. *Journal of microscopy* **224**, 213-232
101. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., and Sternberg, M. J. (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nature protocols* **10**, 845-858

102. Ashkenazy, H., Abadi, S., Martz, E., Chay, O., Mayrose, I., Pupko, T., and Ben-Tal, N. (2016) ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic acids research* **44**, W344-350
103. Wu, C., Orozco, C., Boyer, J., Leglise, M., Goodale, J., Batalov, S., Hodge, C. L., Haase, J., Janes, J., Huss, J. W., 3rd, and Su, A. I. (2009) BioGPS: an extensible and customizable portal for querying and organizing gene annotation resources. *Genome biology* **10**, R130
104. Sierecki, E., Stevers, L. M., Giles, N., Polinkovsky, M. E., Moustaqil, M., Mureev, S., Johnston, W. A., Dahmer-Heath, M., Skalamera, D., Gonda, T. J., Gabrielli, B., Collins, B. M., Alexandrov, K., and Gambin, Y. (2014) Rapid mapping of interactions between Human SNX-BAR proteins measured in vitro by AlphaScreen and single-molecule spectroscopy. *Molecular & cellular proteomics : MCP* **13**, 2233-2245
105. Wassmer, T., Attar, N., Harterink, M., van Weering, J. R., Traer, C. J., Oakley, J., Goud, B., Stephens, D. J., Verkade, P., Korswagen, H. C., and Cullen, P. J. (2009) The retromer coat complex coordinates endosomal sorting and dynein-mediated transport, with carrier recognition by the trans-Golgi network. *Developmental cell* **17**, 110-122
106. Bujny, M. V., Ewels, P. A., Humphrey, S., Attar, N., Jepson, M. A., and Cullen, P. J. (2008) Sorting nexin-1 defines an early phase of Salmonella-containing vacuole-remodeling during Salmonella infection. *J Cell Sci* **121**, 2027-2036
107. Yin, P., Hong, Z., Yang, X., Chung, R. T., and Zhang, L. (2016) A role for retromer in hepatitis C virus replication. *Cellular and molecular life sciences : CMLS* **73**, 869-881
108. Popa, A., Zhang, W., Harrison, M. S., Goodner, K., Kazakov, T., Goodwin, E. C., Lipovsky, A., Burd, C. G., and DiMaio, D. (2015) Direct binding of retromer to human papillomavirus type 16 minor capsid protein L2 mediates endosome exit during viral infection. *PLoS Pathog* **11**, e1004699
109. Ganti, K., Massimi, P., Manzo-Merino, J., Tomaic, V., Pim, D., Playford, M. P., Lizano, M., Roberts, S., Kranjec, C., Doorbar, J., and Banks, L. (2016) Interaction of the Human Papillomavirus E6 Oncoprotein with Sorting Nexin 27 Modulates Endocytic Cargo Transport Pathways. *PLoS Pathog* **12**, e1005854
110. Finsel, I., Ragaz, C., Hoffmann, C., Harrison, C. F., Weber, S., van Rahden, V. A., Johannes, L., and Hilbi, H. (2013) The Legionella effector RidL inhibits retrograde trafficking to promote intracellular replication. *Cell Host Microbe* **14**, 38-50

111. Cozier, G. E., Carlton, J., McGregor, A. H., Gleeson, P. A., Teasdale, R. D., Mellor, H., and Cullen, P. J. (2002) The phox homology (PX) domain-dependent, 3-phosphoinositide-mediated association of sorting nexin-1 with an early sorting endosomal compartment is required for its ability to regulate epidermal growth factor receptor degradation. *J Biol Chem* **277**, 48730-48736
112. Zhong, Q., Watson, M. J., Lazar, C. S., Hounslow, A. M., Waltho, J. P., and Gill, G. N. (2005) Determinants of the endosomal localization of sorting nexin 1. *Molecular biology of the cell* **16**, 2049-2057
113. Paul, B., Kim, H. S., Kerr, M. C., Huston, W. M., Teasdale, R. D., and Collins, B. M. (2017) Structural basis for the hijacking of endosomal sorting nexin proteins by Chlamydia trachomatis. *Elife* **6**
114. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-2132
115. Simonetti, B., Danson, C. M., Heesom, K. J., and Cullen, P. J. (2017) Sequence-dependent cargo recognition by SNX-BARs mediates retromer-independent transport of CI-MPR. *J Cell Biol* **216**, 3695-3712
116. Kvainickas, A., Jimenez-Orgaz, A., Nagele, H., Hu, Z., Dengjel, J., and Steinberg, F. (2017) Cargo-selective SNX-BAR proteins mediate retromer trimer independent retrograde transport. *J Cell Biol* **216**, 3677-3693
117. Seaman, M. N., McCaffery, J. M., and Emr, S. D. (1998) A membrane coat complex essential for endosome-to-Golgi retrograde transport in yeast. *J Cell Biol* **142**, 665-681
118. Simunovic, M., and Voth, G. A. (2015) Membrane tension controls the assembly of curvature-generating proteins. *Nat Commun* **6**, 7219
119. Arighi, C. N., Hartnell, L. M., Aguilar, R. C., Haft, C. R., and Bonifacino, J. S. (2004) Role of the mammalian retromer in sorting of the cation-independent mannose 6-phosphate receptor. *J Cell Biol* **165**, 123-133
120. Seaman, M. N. (2004) Cargo-selective endosomal sorting for retrieval to the Golgi requires retromer. *J Cell Biol* **165**, 111-122
121. Lombardi, D., Soldati, T., Riederer, M. A., Goda, Y., Zerial, M., and Pfeffer, S. R. (1993) Rab9 functions in transport between late endosomes and the trans Golgi network. *EMBO J* **12**, 677-682
122. Meyer, C., Zizioli, D., Lausmann, S., Eskelinen, E. L., Hamann, J., Saftig, P., von Figura, K., and Schu, P. (2000) mu1A-adaptin-deficient mice: lethality, loss of AP-1 binding and rerouting of mannose 6-phosphate receptors. *EMBO J* **19**, 2193-2203

123. Seaman, M. N. J. (2018) Retromer and the cation-independent mannose 6-phosphate receptor-Time for a trial separation? *Traffic* **19**, 150-152
124. Elwell, C. A., Czudnochowski, N., von Dollen, J., Johnson, J. R., Nakagawa, R., Mirrashidi, K., Krogan, N. J., Engel, J. N., and Rosenberg, O. S. (2017) Chlamydia interfere with an interaction between the mannose-6-phosphate receptor and sorting nexins to counteract host restriction. *Elife* **6**
125. Lucas, M., Gershlick, D. C., Vidaurrazaga, A., Rojas, A. L., Bonifacino, J. S., and Hierro, A. (2016) Structural Mechanism for Cargo Recognition by the Retromer Complex. *Cell* **167**, 1623-1635 e1614
126. Harterink, M., Port, F., Lorenowicz, M. J., McGough, I. J., Silhankova, M., Betist, M. C., van Weering, J. R. T., van Heesbeen, R., Middelkoop, T. C., Basler, K., Cullen, P. J., and Korswagen, H. C. (2011) A SNX3-dependent retromer pathway mediates retrograde transport of the Wnt sorting receptor Wntless and is required for Wnt secretion. *Nat Cell Biol* **13**, 914-923
127. Gallon, M., Clairfeuille, T., Steinberg, F., Mas, C., Ghai, R., Sessions, R. B., Teasdale, R. D., Collins, B. M., and Cullen, P. J. (2014) A unique PDZ domain and arrestin-like fold interaction reveals mechanistic details of endocytic recycling by SNX27-retromer. *Proc Natl Acad Sci U S A* **111**, E3604-3613
128. Clairfeuille, T., Mas, C., Chan, A. S., Yang, Z., Tello-Lafoz, M., Chandra, M., Widagdo, J., Kerr, M. C., Paul, B., Merida, I., Teasdale, R. D., Pavlos, N. J., Anggono, V., and Collins, B. M. (2016) A molecular code for endosomal recycling of phosphorylated cargos by the SNX27-retromer complex. *Nat Struct Mol Biol* **23**, 921-932
129. Muhammad, A., Flores, I., Zhang, H., Yu, R., Staniszewski, A., Planel, E., Herman, M., Ho, L., Kreber, R., Honig, L. S., Ganetzky, B., Duff, K., Arancio, O., and Small, S. A. (2008) Retromer deficiency observed in Alzheimer's disease causes hippocampal dysfunction, neurodegeneration, and Abeta accumulation. *Proc Natl Acad Sci U S A* **105**, 7327-7332
130. Vardarajan, B. N., Bruesegem, S. Y., Harbour, M. E., Inzelberg, R., Friedland, R., St George-Hyslop, P., Seaman, M. N., and Farrer, L. A. (2012) Identification of Alzheimer disease-associated variants in genes that regulate retromer function. *Neurobiol Aging* **33**, 2231 e2215-2231 e2230
131. McGeary, R. P., and Fairlie, D. P. (1998) Macrocyclic peptidomimetics: potential for drug development. *Curr Opin Drug Discov Devel* **1**, 208-217

132. Driggers, E. M., Hale, S. P., Lee, J., and Terrett, N. K. (2008) The exploration of macrocycles for drug discovery--an underexploited structural class. *Nat Rev Drug Discov* **7**, 608-624
133. Robinson, J. A. (2011) Protein epitope mimetics as anti-infectives. *Curr Opin Chem Biol* **15**, 379-386
134. Craik, D. J., Fairlie, D. P., Liras, S., and Price, D. (2013) The future of peptide-based drugs. *Chem Biol Drug Des* **81**, 136-147
135. Nowick, J. S. (2008) Exploring beta-sheet structure and interactions with chemical model systems. *Acc Chem Res* **41**, 1319-1330
136. Fasan, R., Dias, R. L., Moehle, K., Zerbe, O., Vrijbloed, J. W., Obrecht, D., and Robinson, J. A. (2004) Using a beta-hairpin to mimic an alpha-helix: cyclic peptidomimetic inhibitors of the p53-HDM2 protein-protein interaction. *Angew Chem Int Ed Engl* **43**, 2109-2112
137. Loughlin, W. A., Tyndall, J. D., Glenn, M. P., and Fairlie, D. P. (2004) Beta-strand mimetics. *Chem Rev* **104**, 6085-6117
138. Moriuchi, T., and Hirao, T. (2004) Highly ordered structures of peptides by using molecular scaffolds. *Chem Soc Rev* **33**, 294-301
139. White, C. J., and Yudin, A. K. (2011) Contemporary strategies for peptide macrocyclization. *Nat Chem* **3**, 509-524
140. Edman, P. (1959) Chemistry of amino acids and peptides. *Annu Rev Biochem* **28**, 69-96
141. Horton, D. A., Bourne, G. T., and Smythe, M. L. (2002) Exploring privileged structures: the combinatorial synthesis of cyclic peptides. *Mol Divers* **5**, 289-304
142. Kwon, Y. U., and Kodadek, T. (2007) Quantitative comparison of the relative cell permeability of cyclic and linear peptides. *Chem Biol* **14**, 671-677
143. Robinson, J. A. (2008) Beta-hairpin peptidomimetics: design, structures and biological activities. *Acc Chem Res* **41**, 1278-1288
144. Milroy, L. G., Grossmann, T. N., Hennig, S., Brunsveld, L., and Ottmann, C. (2014) Modulators of protein-protein interactions. *Chem Rev* **114**, 4695-4748
145. Hill, T. A., Shepherd, N. E., Diness, F., and Fairlie, D. P. (2014) Constraining cyclic peptides to mimic protein structure motifs. *Angew Chem Int Ed Engl* **53**, 13020-13041
146. Nevola, L., and Giralt, E. (2015) Modulating protein-protein interactions: the potential of peptides. *Chem Commun (Camb)* **51**, 3302-3315

147. Berman, D. M., Wilkie, T. M., and Gilman, A. G. (1996) GAIP and RGS4 are GTPase-activating proteins for the Gi subfamily of G protein alpha subunits. *Cell* **86**, 445-452

148. Hepler, J. R., Berman, D. M., Gilman, A. G., and Kozasa, T. (1997) RGS4 and GAIP are GTPase-activating proteins for Gq alpha and block activation of phospholipase C beta by gamma-thio-GTP-Gq alpha. *Proc Natl Acad Sci U S A* **94**, 428-432

149. Longenecker, K. L., Lewis, M. E., Chikumi, H., Gutkind, J. S., and Derewenda, Z. S. (2001) Structure of the RGS-like domain from PDZ-RhoGEF: linking heterotrimeric g protein-coupled signaling to Rho GTPases. *Structure* **9**, 559-569

150. De Vries, L., Zheng, B., Fischer, T., Elenko, E., and Farquhar, M. G. (2000) The regulator of G protein signaling family. *Annu Rev Pharmacol Toxicol* **40**, 235-271

151. Berman, D. M., Kozasa, T., and Gilman, A. G. (1996) The GTPase-activating protein RGS4 stabilizes the transition state for nucleotide hydrolysis. *J Biol Chem* **271**, 27209-27212

152. Chen, C. K., Wieland, T., and Simon, M. I. (1996) RGS-r, a retinal specific RGS protein, binds an intermediate conformation of transducin and enhances recycling. *Proc Natl Acad Sci U S A* **93**, 12885-12889

153. Watson, N., Linder, M. E., Druey, K. M., Kehrl, J. H., and Blumer, K. J. (1996) RGS family members: GTPase-activating proteins for heterotrimeric G-protein alpha-subunits. *Nature* **383**, 172-175

154. Capuano, B., Crosby, I. T., and Lloyd, E. J. (2002) Schizophrenia: genesis, receptorology and current therapeutics. *Curr Med Chem* **9**, 521-548

155. Sadee, W., Hoeg, E., Lucas, J., and Wang, D. (2001) Genetic variations in human G protein-coupled receptors: implications for drug therapy. *AAPS PharmSci* **3**, E22

156. Weiner, D. M., Burstein, E. S., Nash, N., Croston, G. E., Currier, E. A., Vanover, K. E., Harvey, S. C., Donohue, E., Hansen, H. C., Andersson, C. M., Spalding, T. A., Gibson, D. F., Krebs-Thomson, K., Powell, S. B., Geyer, M. A., Hacksell, U., and Brann, M. R. (2001) 5-hydroxytryptamine2A receptor inverse agonists as antipsychotics. *J Pharmacol Exp Ther* **299**, 268-276

157. Zhong, H., and Neubig, R. R. (2001) Regulator of G protein signaling proteins: novel multifunctional drug targets. *J Pharmacol Exp Ther* **297**, 837-845

158. Neubig, R. R., and Siderovski, D. P. (2002) Regulators of G-protein signalling as new central nervous system drug targets. *Nat Rev Drug Discov* **1**, 187-197

159. Abramow-Newerly, M., Roy, A. A., Nunn, C., and Chidiac, P. (2006) RGS proteins have a signalling complex: interactions between RGS proteins and GPCRs, effectors, and auxiliary proteins. *Cell Signal* **18**, 579-591
160. Hollinger, S., and Hepler, J. R. (2002) Cellular regulation of RGS proteins: modulators and integrators of G protein signaling. *Pharmacol Rev* **54**, 527-559
161. Cunningham, M. L., Waldo, G. L., Hollinger, S., Hepler, J. R., and Harden, T. K. (2001) Protein kinase C phosphorylates RGS2 and modulates its capacity for negative regulation of Galpha 11 signaling. *J Biol Chem* **276**, 5438-5444
162. Ogier-Denis, E., Pattingre, S., El Benna, J., and Codogno, P. (2000) Erk1/2-dependent phosphorylation of Galpha-interacting protein stimulates its GTPase accelerating activity and autophagy in human colon cancer cells. *J Biol Chem* **275**, 39090-39095
163. Tu, Y., Popov, S., Slaughter, C., and Ross, E. M. (1999) Palmitoylation of a conserved cysteine in the regulator of G protein signaling (RGS) domain modulates the GTPase-activating activity of RGS4 and RGS10. *J Biol Chem* **274**, 38260-38267
164. Tesmer, J. J., Sunahara, R. K., Gilman, A. G., and Sprang, S. R. (1997) Crystal structure of the catalytic domains of adenylyl cyclase in a complex with Gsalpha.GTPgammaS. *Science* **278**, 1907-1916
165. Hollinger, S., Ramineni, S., and Hepler, J. R. (2003) Phosphorylation of RGS14 by protein kinase A potentiates its activity toward G alpha i. *Biochemistry* **42**, 811-819
166. Lin, C., Koval, A., Tishchenko, S., Gabdulkhakov, A., Tin, U., Solis, G. P., and Katanaev, V. L. (2014) Double suppression of the Galpha protein activity by RGS proteins. *Mol Cell* **53**, 663-671
167. Martin, S., Okano, S., Kistler, C., Fernandez-Rojo, M. A., Hill, M. M., and Parton, R. G. (2009) Spatiotemporal regulation of early lipolytic signaling in adipocytes. *J Biol Chem* **284**, 32097-32107
168. Wang, C. W., Miao, Y. H., and Chang, Y. S. (2014) A sterol-enriched vacuolar microdomain mediates stationary phase lipophagy in budding yeast. *J Cell Biol* **206**, 357-366
169. de Marco, A. (2007) Protocol for preparing proteins with improved solubility by co-expressing with molecular chaperones in Escherichia coli. *Nat Protoc* **2**, 2632-2639
170. Mesmin, B., Bigay, J., Moser von Filseck, J., Lacas-Gervais, S., Drin, G., and Antonny, B. (2013) A four-step cycle driven by PI(4)P hydrolysis directs sterol/PI(4)P exchange by the ER-Golgi tether OSBP. *Cell* **155**, 830-843